CHAPTER 1

The Assessment of Individuals

The Critical Role and Fundamentals of Measurement

The importance of measurement in our daily lives and in research in education and the social sciences cannot be overstated. How well a construct is measured is critical in so many different ways. Consider the importance of measuring the height and weight of a newborn baby. These are general indicators of the health of the baby. If a measurement is unusual, then actions are taken to bring the measurement more in line with what is considered typical. Consider any college or university course taken. Achievement tests to determine how much students have mastered the course content are the norm. If the test is flawed in some way, this may have a negative impact on GPA, which would have further consequences. Consider tests of ability that are used for streaming junior high school students into a university-bound or non-university-bound set of courses in high school. Based on the test score, a student's life is affected. Consider a job interview where a panel of judges rates applicant responses to interview questions. Based at least partly on their ratings (measurement of applicant performance in the interview), a job offer may or may not be forthcoming. Consider how carefully politicians pay attention to popularity polls. Their future careers rest on how this information is collected and portrayed.

On the research side of things, if the measures used in the study that is being carried out are questionable, the research is not going to be published. If a poor measure of job satisfaction is used, then the likelihood of it being related to other variables of interest to the researcher (such as intentions to quit the organization) is also poor; the analyses are less likely to be able to detect the relationships hypothesized. The measure that is being used in research should exactly measure the construct of interest. For example, in a measure of job satisfaction, there may be a couple of items that actually measure knowledge of organizational policies. If this is the case, then that measure is impure or contaminated.

Measurement is used all the time in our daily lives and it is an integral part of the research process. Knowledge about measurement—how to correctly assess constructs, how to critically examine others' use of measures, and how to be a smart consumer of published tests—is an important skill of a social scientist. This book is written for that reason. At the end, you should know how to construct a test, how to evaluate a test, and how much faith you can put in the scores of any given instrument.

Measurement in the Physical Sciences

Those of us in the social sciences are often envious of the precision with which physical scientists are able to measure their constructs. There is not a lot of quarreling about the temperature, speed, height, weight, luminance level, or color of a given substance or event. The instruments that have been designed to measure such constructs have been built to be reliable and are usually calibrated on a regular basis to ensure the accuracy of the values that they produce. As long as the individual using the instrument knows how to use it and knows how to interpret the values, there is no problem in the measurement aspect of the work.

Measurement in the Social Sciences

On the other hand, social scientists are often dealing with ambiguous constructs such as political activism, delinquency, leadership, intelligence, personality, creativity, depression, anxiety, and so forth. Not only is there disagreement on how these are measured but also, in many cases, there is no overall agreement on what is meant by the construct itself. Thus, social scientists battle on two fronts. The first thing to do when preparing to develop or use a test is to be absolutely clear about what it is that you want to measure. This is called the conceptual definition of the construct. For example, if you want to measure creativity, you must first define for yourself, and therefore for all who will read your work, what you mean by creativity. As a creativity test consumer, you will first want to determine how much you agree with how the test author defined creativity. If you don't agree, then don't purchase the test.

Even after leaping the first hurdle of getting an audience to agree with a conceptual definition, social scientists must then convince them that how that construct is measured is an accurate representation of the construct. That is, translating the conceptual definition into an operational definition (in measurement, this usually means creating items to assess the construct) requires careful and methodical work. Two chapters are devoted to this exercise—one to creating items and the other to creating responses to those items.

Thus, measurement in the social sciences is fraught with pitfalls and yet is such a critical skill that it is well worth cultivating. Before moving on to introduce the topic of construct definition, a review of some of the highlights in the history of individual difference measurement, or, more technically, psychometrics, is presented.

Historical Highlights of Measurement

Assessment of individual differences has a very long history. The Chinese civil service in 2200 B.C. was the first recorded group to use formal testing procedures for selection and performance appraisal (Bowman, 1989). This system was the model for British, French, and German governments in the 19th century. The ancient Greeks also used individual difference testing (Doyle, 1974).

Measurement and testing, however, received a great boost in the 19th century due to the rising interest in several countries about various aspects of individual differences. The controversial and revolutionary writings of the English naturalist Charles Darwin; the work in the experimental psychology laboratories of Wundt, Ebbinghaus, and Fechner in Germany; the study of intelligence in France by Binet and Simon; the work of English biologists Galton and Pearson; and the American experimental psychologist Cattell all contributed in tangential or direct ways to the testing movement.

A seminal event in testing occurred when Alfred Binet, a French psychologist, and Theodore Simon were commissioned by the Parisian minister of public education in 1904 to develop a process to identify schoolchildren who would not benefit from instruction in the regular school system. Their work on the first formal intelligence test resulted in the assessment of children aged 3 to 13.

Work on other tests of intelligence, achievement, personality, and interests took place in the early 20th century. The advent of the First World War, and the need to test intelligence for large numbers of people in group settings, rendered the individually administered tests that had been developed to date too resource intensive. The result was the development by Otis in 1917 of the Army Alpha (for literate respondents) and Army Beta (for illiterate respondents) group-administered intelligence tests. The current Armed Services Vocational Aptitude Battery is based on Otis's early work.

The need for matching the vocational skills and interests of the many new immigrants to North America was answered by the development of interest inventories. As the standard of living for many living in North America climbed, more young adults wanted to enter colleges, universities, and graduate schools. The need for tests of achievement that assessed students and allowed for comparison with others across the continent (i.e., standardized testing) influenced the creation and use of the Scholastic Aptitude Test, the Graduate Record Examination, and many others.

It was during World War II that, for the first time, the capabilities of machines outpaced the capabilities of humans. The need to develop careful tests of psychomotor skills was answered. In the 1930s and 1940s, the interest in personality as a construct was widely discussed, with Freud's and Jung's writings capturing the imaginations of laypeople and professionals alike. Today, measures of various aspects of personality are commonplace.

Testing has become normative in schools and workplaces. However, some identifiable demographic groups have been disadvantaged by traditional paper-and-pencil tests. With the civil rights movement and the passage of Title VII in the United States, the testing enterprise went on the defensive. Specifically, the onus of the "validity of a test" was on the test administrator. Litigation abounded and many organizations became wary of using test scores for making personnel decisions.

The 1980s saw an exponential rise in the use of computers. Computer-based testing and computer-adaptive testing have become more regular features of the testing terrain. New tests are coming on the scene every day, revisions to older tests recur on a regular basis, and the public is increasingly knowledgeable about tests and their rights with regard to the use of test scores. All of these advances testify to the need for social scientists to be skilled in test development and evaluation methods.

Statistics Background

Before continuing, it will be useful to freshen up on some statistics basics. As this book proceeds, some fairly advanced statistical information will be introduced. This book assumes that the reader has taken (and passed) a basic statistics course in college or university. Topics that will be reviewed here are scales of measurement, characteristics of the normal distribution, *p* values, and statistical significance. In addition, a quick refresher on sampling distributions, correlation, and regression is in order. Finally, linear conversion of raw scores is presented, as this is used extensively in the measurement literature.

Scales of Measurement. In the measurement process, data are collected and numbers assigned to them. Depending on the type of data collected, those numbers carry different meanings. These meanings are based on the scale of measurement used.

The most rudimentary scale of measurement is the nominal scale. Here, the numbers are simply ways to code categorical information. For example, data may be collected on men and women and, for sake of expediency, all cases that are men are coded with a 1 and all cases that are women are coded with a 2. If data are collected on a college campus, college major may be coded numerically (e.g., science = 1, social science = 2, humanities = 3, etc.). In all instances, these nominal numbers reflect nothing other than a category. The numerical values in nominal scores do not represent an assessment of more or less of any particular value.

The next, more complex, level of measurement is ordinal. In ordinal measurement, the numbers assigned have meaning in that they demonstrate a rank order of the cases. For example, if members of a class are rank ordered from highest to lowest based on their test scores, the rank ordering indicates who did better than whom. However, ordinal measures do not indicate an absolute level of performance. For example, if the finishers of a race are rank ordered from first to last, this indicates who ran faster than whom but does not indicate anything about the runners' absolute speed in completing the race.

Interval is the next most sophisticated level of measurement. Here, the numbers are rank ordered, but now more information is contained in the numbers. Specifically, the differences between the numbers are equivalent. That is, the difference between 2 and 3 can be assumed to be the same as the difference between 3 and 4. For example, temperature is measured at the interval level. If it is 20 degrees on Day 1, 25 degrees on Day 2, and 30 degrees on Day 3, the temperature change from Day 1 to Day 2 is exactly the same as that from Day 2 to Day 3.

For interval-level data, each case has an absolute value associated with it. However, there is no fixed zero point with these types of scales. The result of no fixed zero is best demonstrated through an example. Let's say we want to measure individuals' "need for achievement" with a particular scale. With an interval level of measurement, the scores can be interpreted to mean that someone with a score of 15 is as different from someone with a score of 20 as is someone with a score of 20 compared to someone with a score of 25. The scale provides us with interval-level information. However, there is no universally accepted level of "zero need for achievement." Therefore, we cannot conclude that someone with a score of 20 has half the need for achievement as does a person with a score of 40. In order to make that claim, we would need to have an absolute zero point on the scale.

Ratio level of measurement provides the most information about the numbers because it has all the characteristics of interval-level measurement, plus there is an absolute zero point. Scales measured at the ratio level would include height, weight, speed, time, and distance. If person A is six feet tall and person B is three feet tall, it is true to say that person A is twice as tall as person B. If person A runs 10 kilometers in 40 minutes and person B runs 5 kilometers in 40 minutes, it is true to say that person B ran half as fast as person A.

The reason for the review of this topic is that the appropriate statistical procedure to use in any data set depends on the level of measurement used. Most data that social scientists collect are at the nominal, ordinal, or interval level. In scale development and use, we often aspire to measure at the interval level, but we can often only achieve the ordinal level.

The Normal Distribution. A common assumption about any measured individual difference, whether it is a personality characteristic, cognitive skill, motor skill, social skill, or other attribute, is that this difference is normally distributed in the population. The normal distribution is a symmetrical, bell-shaped curve (see Figure 1.1). The shape shows that more of the area under the curve is in the center of the distribution, and, as one moves toward the "tails" of the distribution, the area under the curve becomes less and less.

Using height as an example of a normally distributed characteristic, everyone's height in a country could be measured. It would be the case that there are a few short people and a few tall people, but most people's heights would fall somewhere in the midrange. The more extreme the height (shorter or taller), the fewer the number of people who would have that height. The normal distribution serves to determine if a particular value is extreme or not when conducting statistical analyses. Values at the extreme ends of the distribution are unusual and the exact "extremeness" of any value can be quantified based on probability, which we turn to next.

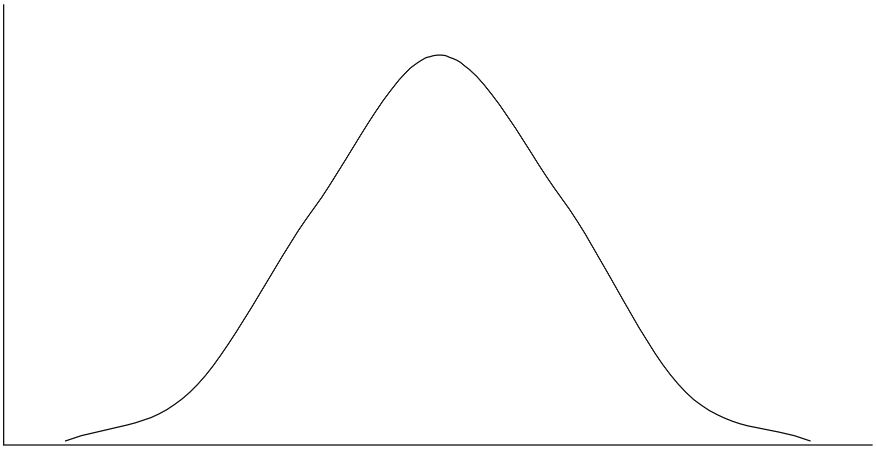

Figure 1.1 Normal Distribution

Probability and Statistical Significance. Prior to beginning this section, recall that when probability is mentioned, science is a conservative endeavor. This means that when scientists ask a question such as, Are girls more likely than boys to sign up for a high school auto mechanics class? they are likely to say that the question is empirical and that data should be collected to answer the question.

So a sample of high schools in the city is selected, and the percentages of girls and boys enrolled in auto mechanics classes are compared. Suppose it is found that, out of the 1,000 enrollments in the high school auto mechanics classes for the fall term, 55% were boys and 45% were girls. Would it be justifiable to claim that more boys than girls signed up for auto mechanics? What if the percentages were 60% to 40%? What if the percentages were 75% to 25%? What if the percentages were 90% to 10%? At what point would the scientists be willing to say that there is a "statistically significant difference in the proportion of boys versus girls taking auto mechanics"?

The answer is known as the adopted *alpha level* (α). It reports that the difference found in the sample of boys versus girls would happen by chance alone *X* number of times out of 100. So what does *X* equal? Usually it equals 1 or 5. This means that the difference in percentages found in the sample would have to be large enough to only occur by chance 1 out of 100 times ($\alpha = 0.01$); or, less conservatively, 5 out of 100 times ($\alpha = 0.05$); or, even less conservatively, 10 out of 100 times ($\alpha = 0.10$).

These α levels correspond to *p values,* or sometimes *p levels,* on statistical printouts. The *p* value stands for the probability level. If the *p* value for a particular statistical test (whether it is a correlation, *t* test, chi-square, etc.) is equal to 0.03, then this is interpreted to mean that the finding from the particular sample would occur by chance alone 3 times out of 100. If the *p* value was equal to 0.30, then this is interpreted to mean that the finding from the particular sample would occur by chance alone 30 times out of 100. If the *p* value was equal to 0.006, then this is interpreted to mean that the finding from the particular sample would occur by chance alone 6 times out of 1,000. In the social sciences, the usual α level adopted for making decisions about statistical significance is 0.05 or 0.01.

Sampling Distributions. There is a difference between sample distributions and sampling distributions. An example of a sample distribution would be a distribution of a set of scores on a history test by a class of 6th-grade students. This distribution would show the mean score of the class, variance of the class scores, lowest and highest scores, and so forth. Sampling distributions, however, are theoretical distributions and are used in making statistical decisions. Like sample distributions, sampling distributions have means and variances. Multiple sampling distributions are associated with inferential statistics, such as *t* tests, *F* tests, chi-square tests, correlation tests, and so forth. The shape of each sampling distribution is based on different sample sizes and the number of variables in the analysis. Sampling distributions are used to set the α level for a particular statistical test and used to decide whether or not to reject the null hypothesis.

For example, if we were interested in the difference between need for achievement scores for men and women and we had a sample of 10 men and 10 women, we would test for the differences between the means of the sample scores and have a *t* statistic generated. We would then use a *t* table that reports the critical value the calculated *t* value needs to exceed in order for it be considered *statistically significant.* That is, the *t* value calculated has to be extreme enough to be considered highly unusual and not likely to occur by chance alone.

Sample sizes are important because they tell which sampling distribution to use to test whether or not the calculated statistic is significant or not. What is common about all sampling distributions is that as the sample size on which the statistic is calculated increases, the *critical value* the statistic needs to exceed to be considered significant (i.e., extreme) decreases. Take, for example, our 10 men and 10 women and their hypothesized difference in need for achievement scores. If we had used a two-tailed test and adopted an α level of 0.05, then the critical *t* value the calculated *t* value needs to exceed is 2.101. All else remaining constant, if we had 15 men and 15 women in our sample, the *t* value to exceed is 2.048. Thus, it is easier to find a significant difference using the larger sample than with the smaller sample.

Correlation. Correlation describes the strength and direction of the linear relationship between two variables. Data for a correlational analysis are put into two columns (vectors) of numbers, where *X* represents values on one variable and *Y* represents values on the other variable. These columns would be set up like the following:

X	*Y*

If the *X* variable was the number of hours studied, the *Y* variable might represent grades on an exam on that material as follows:

X	*Y*
5	80
6	87
7	89
9	95

A general pattern in the four pairs of scores emerges: as the number of hours of study goes up, the grade on the exam goes up. That is, the pairs vary together in a linear, positive manner. Let's take another example. What if the *X* variable was a measure of job satisfaction (where higher scores mean higher levels of satisfaction) and *Y* was a measure of intentions to quit? Then the pairs of numbers might look like the following:

X	*Y*
10	3
8	5
7	7
2	10

In this example, a general pattern in the four pairs of scores also emerges. However, this time, as the job satisfaction values go down, intentions to quit values go up. So in this example, the pairs vary together in a linear, negative manner.

The statistic that summarizes the strength and direction of the relationship between two vectors of variables is called the *Pearson product-moment correlation coefficient,* or *correlation coefficient* for short. Values of the correlation coefficient vary from −1.00 to +1.00. The more the pairs of values vary together, the stronger the relationship and the farther from 0.00 (whether a positive or negative value) the correlation coefficient will be. That is, a correlation coefficient of −0.80 indicates that there is a strong negative relationship between the pairs of values. A correlation coefficient of 0.40 indicates that there is a moderate positive relationship between the pairs of values.

Table 1.1 shows a set of four scores: *A, B, X,* and *Y.* In Box 1.1, the correlation between *A* and *B* is calculated to review the procedure. However, given the common availability of many of these calculations in computer programs, this book takes the approach that interpreting the information on the outputs provided by such programs is worthy of discussion. Therefore, the correlation program in SPSS will be used to first assess the correlation between *A* and *B,* and then again between *X* and *Y.* The relevant sections of the printout are shown in Box 1.2.

Table 1.1 Data for Two Examples of Pearson Correlations

Case	*A*	*B*	*X*	*Y*
1	6	45	10	1
2	7	120	9	3
3	8	100	8	2
4	9	101	7	4
5	2	76	6	3
6	3	55	5	5
7	4	80	4	3
8	5	76	3	4
9	6	90	2	5
10	7	110	1	5
11	8	115	10	6
12	9	120	9	3
13	1	52	8	2
14	2	40	7	4
15	3	43	6	1
16	4	20	5	6
17	5	86	4	5
18	5	80	3	4
19	6	15	2	1
20	7	87	1	2

Box 1.1 Computation of the Pearson Correlation of Columns *A* and *B* in Table 1.1

The Pearson Correlation for two variables (bivariate) computational formula is as follows:

$$r = [N\Sigma AB - (\Sigma A)\ (\Sigma B)]/\sqrt{[N\ (\Sigma A^2) - (\Sigma A)^2]\ [N\ (\Sigma B^2) - (\Sigma B)^2]} \qquad (1–1)$$

This formula is used to calculate the correlation of the data in columns *A* and *B* in Table 1.1. First, the values in columns *A* and *B* are squared and then *A* and *B* are cross-multiplied. The results are shown in Table 1.2.

(Continued)

Box 1.1 (Continued)

Table 1.2 Bivariate Pearson Correlation Computation Example Data Set

Case	*A*	*B*	*A*^{2}	*B*^{2}	*AB*
1	6	45	36	2025	270
2	7	120	49	14400	840
3	8	100	64	10000	800
4	9	101	81	10201	909
5	2	76	4	5776	152
6	3	55	9	3025	165
7	4	80	16	6400	320
8	5	76	25	5776	380
9	6	90	36	8100	540
10	7	110	49	12100	770
11	8	115	64	13225	920
12	9	120	81	14400	1080
13	1	52	1	2704	52
14	2	40	4	1600	80
15	3	43	9	1849	129
16	4	20	16	400	80
17	5	86	25	7396	430
18	5	80	25	6400	400
19	6	15	36	225	90
20	7	87	49	7569	609
SUM (Σ)	107	1511	679	133571	9016

$$r = [(20 \times 9016) - (107)\ (1511)]/\sqrt{[20(679) - (107)^2]\ [20(133571) - (1511)^2]}$$

$$= (180320 - 161677)/\sqrt{[(13580) - (11449)]\ [(2671420) - (2283121)]}$$

$$= (18643)/\sqrt{(2131)\ (383299)}$$

$$= 18643/28766$$

$$= 0.684$$

The results from the printout show that the correlation coefficient calculated between *A* and *B* is equal to 0.648. The significance level (or *p* value) of 0.002 indicates that the chance of us finding the magnitude of relationship between these

Box 1.2 Output from the SPSS Pearson Bivariate Correlational Analyses of Columns *A* and *B* and Columns *X* and *Y* in Table 1.1

Pearson Correlation between *A* and *B*: 0.648[a]
Sig. (two-tailed): 0.002; $N = 20$
a. Correlation is significant ($p < 0.01$; two-tailed)

Pearson Correlation between *X* and *Y*: –0.150[a]
Sig. (two-tailed): 0.527; $N = 20$
a. Correlation is not significant ($p > 0.05$; two-tailed)

20 pairs of numbers by chance alone is 2 in 1000 times. This is even less common (i.e., the finding of 0.648 is a value that would be found at the very extreme upper end of the sampling distribution) than the usual threshold of 0.01 or 0.05, and so it can be concluded that there is a significant positive relationship between *A* and *B*. The two-tailed test is the default for SPSS for testing correlation coefficients. This means that the direction of the relationship between *A* and *B* was not specified in advance. If the direction was specified to be either positive or negative, the option "one-tailed" in the SPSS program could be selected.

The correlation coefficient calculated for *X* and *Y* is equal to –0.150. It is negative in value and has a significance level of 0.527. This indicates that the chances of finding the calculated magnitude of relationship between these 20 pairs of numbers by chance is 527 in 1000 times. This is much higher than the usual threshold of 0.01 or 0.05 (i.e., the finding of –0.150 is a value that would be found in the middle of the sampling distribution), and so it must be concluded that there is no relationship between *X* and *Y*. That is, unless the significance level is smaller than 0.05 or 0.01 (whichever is adopted), it is assumed that the calculated value is not significantly different from 0.00 (the value at the exact middle of the sampling distribution).

Sometimes the question is raised as to how many cases one needs to calculate a correlation coefficient. The answer is, at a bare minimum, three. This is because there need to be at least three cases for the correlation significance level to be calculated. However, three cases are hardly enough to be confident about the calculated value of the correlation coefficient. Exactly how many cases are needed can be directly assessed, depending on the strength of the expected relationship, through something called a power analysis (e.g., Cohen, 1988). However, a good rough rule of thumb is to have 10 cases per variable. Because correlation uses two variables, 20 cases are usually sufficient to be able to draw some conclusions about the sample data if a moderate relationship between them is expected.

Linear Regression. Regression, as in correlation, is an analysis of linear relationships between variables. However, a major difference is that regression requires the researcher to indicate that one variable is dependent (criterion) on the other(s) (predictor[s]). In the linear regression examples in this book, there will always be only one criterion variable. In some cases there will be one predictor (simple or bivariate regression) and in other cases more than one predictor (multiple regression).

A regression analysis produces a series of results that will take a bit of time to review. As in correlation, assume that there are vectors (or columns) of numbers: each column represents a variable and each row represents a case, or a subject's scores on each of the variables. First, a case of simple regression will be reviewed, where there is only one predictor, followed by an example of multiple regression, with two predictors.

Assume the criterion is "starting salary" in $1000 dollars and the predictor is "university GPA." Another predictor, "cognitive ability," will be added in the second analysis. The data are set up as in Table 1.3 and the computational calculations of the bivariate regression are presented in Box 1.3. The SPSS output for the bivariate analysis is shown in Box 1.4.

Text continues on page 17

Table 1.3 Data for Bivariate (Simple) and Multiple Linear Regression Analyses

	Starting Salary ($1,000)	*University GPA*	*Cognitive Ability*
Case 1	20	2.0	100
Case 2	21	2.1	120
Case 3	22	2.0	110
Case 4	23	2.3	104
Case 5	24	2.1	90
Case 6	25	3.5	95
Case 7	26	3.0	115
Case 8	27	2.9	112
Case 9	28	3.4	115
Case 10	29	2.8	98
Case 11	30	3.0	120
Case 12	31	3.3	100
Case 13	32	3.4	110
Case 14	33	2.9	115
Case 15	34	2.8	100
Case 16	35	3.5	102
Case 17	36	3.4	108
Case 18	37	3.3	110
Case 19	38	3.2	116
Case 20	39	3.0	118

Simple Regression Output:

Model Summary			
R	*R-Square*	*Adjusted R-Square*	*Std. Error of Estimate*
0.70	0.49	0.46	4.3378

ANOVA					
	Sums of Squares	*df*	*Mean Square*	*F*	*Sig.*
Regression	326.23	1	326.30	17.34	0.001
Residual	338.70	18	18.82		
Total	665.00	19			

Coefficients					
Model	*Unstandardized b*	*Standard Error*	*Standardized (Beta)*	*t*	*Sig.*
Constant	6.54	5.60		1.17	0.258
GPA	7.93	1.90	0.70	4.16	0.001

Multiple Regression Output:

Model Summary			
R	*R-Square*	*Adjusted R-Square*	*Std. Error of Estimate*
0.72	0.52	0.46	4.3559

ANOVA					
	Sums of Squares	*df*	*Mean Square*	*F*	*Sig.*
Regression	342.44	2	171.30	9.02	0.002
Residual	322.56	17	18.97		
Total	665.00	19			

Coefficients					
Model	*Unstandardized b*	*Standard Error*	*Standardized (Beta)*	*t*	*Sig.*
Constant	–4.25	12.99		–0.33	0.747
GPA	7.72	1.93	0.68	4.01	0.001
Cognitive	0.11	0.12	0.16	0.92	0.369

Box 1.3 Computational Calculations of a Bivariate Regression Analysis

Using the data in Table 1.3, computational formulae are used to generate the regression line and other statistics for the bivariate regression of salary on GPA (see Table 1.4).

The predicted salary scores (using the regression line formula calculated), and squared deviation scores can then be calculated. These are needed for the calculation of the R^2 and standard error of estimate.

Presented next are the predictor scores, the predictor scores less the mean of the predictor scores, and the squares of the difference of those terms. These are needed for the calculation of the standard error of the regression coefficient (*b*).

Table 1.4 Data for Use in Calculating the Bivariate Regression Line

Case	*Salary*	*GPA*	*GPA*^{2}	*GPA* × *Salary*
1	20	2	4	40
2	21	2.1	4.41	44.1
3	22	2	4	44
4	23	2.3	5.29	52.9
5	24	2.1	4.41	50.4
6	25	3.5	12.25	87.5
7	26	3	9	78
8	27	2.9	8.41	78.3
9	28	3.4	11.56	95.2
10	29	2.8	7.84	81.2
11	30	3	9	90
12	31	3.3	10.89	102.3
13	32	3.4	11.56	108.8
14	33	2.9	8.41	95.7
15	34	2.8	7.84	95.2
16	35	3.5	12.25	122.5
17	36	3.4	11.56	122.4
18	37	3.3	10.89	122.1
19	38	3.2	10.24	121.6
20	39	3	9	117
Sum (Σ)	590	57.9	172.81	1749.2
Mean	29.5	2.895		

Predicted Salary	*(Actual – Predicted)*^{2}	*(Actual –* $\bar{Y}$*)*^{2}	*(Predicted –* $\bar{Y}$*)*^{2}
22.4	5.76	90.25	50.41
23.193	4.809249	72.25	39.778249
22.4	0.16	56.25	50.41
24.779	3.164841	42.25	22.287841
23.193	0.651249	30.25	39.778249
34.295	86.397025	20.25	22.992025
30.33	18.7489	12.25	0.6889
29.537	6.436369	6.25	0.001369
33.502	30.272004	2.25	16.016004
28.744	0.065536	0.25	0.571536
30.33	0.1089	0.25	0.6889
32.709	2.920681	2.25	10.297681
33.502	2.256004	6.25	16.016004
29.537	11.992369	12.25	0.001369
28.744	27.625536	20.25	0.571536
34.295	0.497025	30.25	22.992025
33.502	6.240004	42.25	16.016004
32.709	18.412681	56.25	10.297681
31.916	37.015056	72.25	5.837056
30.33	75.1689	90.25	0.6889
	Σ338.70[a]	Σ665	Σ326.34[b]

[a] This is also known as the residual sums of squares.

[b] This is also known as the regression sums of squares. Note there is a slight discrepancy from the printout version due to rounding error in generating the predicted scores.

(Continued)

Box 1.3 (Continued)

GPA	*GPA – Mean GPA*	*(GPA – Mean GPA)*^{2}
2	–0.895	0.801025
2.1	–0.795	0.632025
2	–0.895	0.801025
2.3	–0.595	0.354025
2.1	–0.795	0.632025
3.5	0.605	0.366025
3	0.105	0.011025
2.9	0.005	2.5E-05
3.4	0.505	0.255025
2.8	–0.095	0.009025
3	0.105	0.011025
3.3	0.405	0.164025
3.4	0.505	0.255025
2.9	0.005	2.5E-05
2.8	–0.095	0.009025
3.5	0.605	0.366025
3.4	0.505	0.255025
3.3	0.405	0.164025
3.2	0.305	0.093025
3	0.105	0.011025
$\bar{X} = 2.895$		

Bivariate regression coefficient computational formula:

(1–2) $b = [N \times \Sigma XY - (\Sigma X)(\Sigma Y)]/[N\Sigma X^2 - (\Sigma X)^2]$.

Bivariate regression constant computational formula:

(1–3) $a = \bar{Y} - b\bar{X}$.

To solve for *b,*

$b = [(20)(1749.2) - (57.9)(590)]/[20(172.81) - (57.9)^2]$,
$= (34984 - 34161)/(3456.2 - 3352.41)$,
$= 823/103.79$,
$= 7.93$.

To solve for *a,* then,

$a = \bar{Y} - b\bar{X}$,
$= 29.5 - (7.93)\ (2.895)$,
$= 29.5 - 22.96$,
$= 6.54$.

Regression line: predicted salary = 6.54 + 7.93(GPA)

Calculating the R^2 value:

(1–4) $R^2 = \text{S}\ \Sigma(\bar{Y} - \bar{Y})^2/\Sigma(\bar{Y} - \bar{Y})^2,$
$= 326.34/665,$
$= 0.49.$

Calculating the adjusted R^2 value:

(1–5) Adjusted $R^2 = 1 - (1 - R^2)[(N - 1)/(N - k - 1)],$
$= 1 - (1 - 0.49)\ [(20 - 1)/(20 - 1 - 1)],$
$= 1 - (0.51)(19/18),$
$= 1 - (0.51)(1.06),$
$= 1 - 0.54,$
$= 0.46.$

Calculating the standard error of estimate:

(1–6) $SE = \sqrt{\Sigma(Y - Y')^2/(N - k - 1)},$

where Y = actual scores, Y' = predicted scores, N = sample size, and k = number of predictors,

$SE = \sqrt{338.70/(20 - 1 - 1)},$
$SE = 4.34.$

Calculating the *F:*

(1–7) F = (Regression Sums of Squares/*df*)/(Residual Sums of Squares/*df*),
$= (326.34/1)/(338.70/18),$
$= 326.34/18.82,$
$= 17.34$ (1,18 degrees of freedom).

Calculating the standard error of *b:*

(1–8) $\text{S}b = \sqrt{(SE)^2/(\textit{Sum of Squared Deviations of } X)},$
$= \sqrt{(4.34)^2/5.19},$
$= \sqrt{18.84/5.19},$
$= 1.90.$

Calculating the *t:*

(1–9) $t = b/\text{S}b,$
$= 7.93/1.90,$
$= 4.17$. (This value is the same as that found in the computer printout within rounding error.)

Referring to the information in Box 1.4, there are three tables in the output of an SPSS regression analysis: the model summary, the ANOVA table, and the coefficient table. The model summary and ANOVA tables indicate whether or not *all* of predictors, as a unit, account for a significant amount of variance in the criterion. In the case of simple regression, there is only one predictor, so "all of them as a unit" means only GPA. In the model summary, the *R* value is the multiple correlation between the predictor and criterion. For this example, the value is 0.70 and it is actually calculated

Box 1.4 Bivariate (Simple) Regression Output From SPSS

Model Summary			
R	*R-Square*	*Adjusted R-Square*	*Std. Error of Estimate*
0.70	0.49	0.46	4.3378

ANOVA					
	Sums of Squares	*df*	*Mean Square*	*F*	*Sig.*
Regression	326.23	1	326.30	17.34	0.001
Residual	338.70	18	18.82		
Total	665.00	19			

Coefficients					
Model	*Unstandardized b*	*Standard Error*	*Standardized (Beta)*	*t*	*Sig.*
Constant	6.54	5.60		1.17	0.258
GPA	7.93	1.90	0.70	4.16	0.001

based on the *R*-square value that indicates how much variance in the criterion (salary) can be predicted with the predictor (GPA). In this example, it is 0.49, or 49%.

The adjusted *R*-square value provides an estimate of what to expect the *R*-square value to be if the study was conducted again with a new sample of 20 cases. In this case, the value is 0.46, indicating that some shrinkage in the *R*-square value is expected (i.e., reduced from 0.49 to 0.46). Although the adjusted *R*-square value is always smaller than the *R*-square value, the less shrinkage between the calculated *R*-square and adjusted *R*-square the better. When the difference between them is small, one can be more confident about the robustness of the *R*-square value. Next, the standard error is reported. This is the error associated with predicting scores on the criterion. Larger values indicate more error in prediction than do smaller values.

Whether or not the *R*-square value is significant is determined in the ANOVA table, using an *F* test. It can be seen that the *F* associated with the amount of variance accounted for in starting salary by GPA is 17.34 (with 1 and 18 degrees of freedom), which is significant at 0.001 (i.e., this extreme a value would occur by chance alone only 1 time out of 1,000). Therefore, concluding that university GPA accounts for a significant amount of variance in starting salary is justified.

One of the more important calculations from the analysis is the regression line. The summary of those calculations is in the coefficients table. The regression line is a mathematical function that relates the predictor to the criterion variable. It is calculated so that the line minimizes the squared distances of each point from the line. In simple regression, the regression line is written with the following formula:

$$Y' = a + bX, \tag{1–10}$$

where Y' = the predicted Y score for a given value of X, b = the regression coefficient (also called the *slope*), a = the intercept (or constant), where the regression line crosses the y-axis, and X = the obtained values on X.

To interpret the information in the coefficients table, use the unstandardized constant and GPA coefficients (6.54 and 7.93, respectively). The regression line then can be written as follows:

$$Y' = 6.55 + 7.93(\text{GPA}).$$

For someone with a GPA of 4.0, the predicted starting salary would be

$$Y' = 6.55 + 7.93(4.0),$$

or 36.11 ($36,110, as salary level was coded in $1,000 units).

A scatterplot of the two variables is shown in Figure 1.2 as well as the regression line.

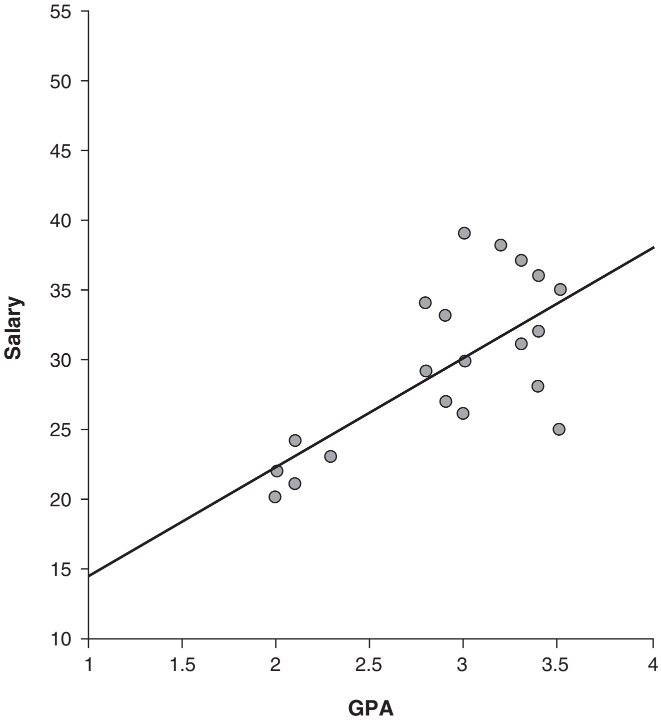

Figure 1.2 Regression Line of Starting Salary on GPA

In addition to obtaining the information for the regression line, the coefficients table reports the "unique" contributions of each predictor and whether the unique contribution is significantly different from zero. If the unstandardized coefficient value is divided by its respective standard error, a *t* value for the predictor is obtained. In this example, there is one predictor (i.e., GPA) and it has a *t* value of 4.16 (7.93/1.90), which is significant at 0.001. When there is only one predictor, the *F* value in the ANOVA table is equal to the square of the *t* value ($4.16^2 = 17.34$). The beta value, or standardized coefficient, is simply the standardized value of the unstandardized coefficient. That is, if standard scores rather than raw scores had been used in this analysis, the beta value would be 0.70. Beta values are like correlations insofar as they range in value from −1.00 to 0.00 to a high of 1.00. A score of 0.00 shows no relationship between the predictor and criterion. Its use will become more apparent in multiple regression, which we will turn to next.

As an example of a multiple regression analysis, cognitive ability scores will be added as a predictor (see Box 1.5). Because cognitive ability is added *after* GPA, a *hierarchical* approach is used in entering the variables into the equation. If both GPA and cognitive ability were entered into the equation at the same time, this would have been a *simultaneous,* or *direct,* entry of the predictors. If the computer selected which variable to enter first into the regression analysis based on a statistical criterion, it would be called a statistically driven entry, the most common of which is *stepwise.* In a stepwise regression analysis, the predictor with the highest zero-order correlation with the criterion is entered into the equation first. Then the predictor with the highest correlation with the criterion, after the effects of the first predictor are taken into account, is entered on the next step. Subsequent steps continue until there are no more variables left that account for a significant amount of the variance in the criterion.

Notice that the *R*-square value in this analysis is 0.52. By adding cognitive ability into the mix, an additional 3% of the variance in starting salary can be accounted for. The overall value of 52% is significant ($F = 9.02$, significant at 0.002), indicating that together GPA and cognitive ability account for a significant amount of variance in starting salary. The adjusted *R*-square is 0.46, indicating that the shrinkage estimate is calculated to be 6% (0.52 − 0.46). This shrinkage is larger than in the previous example with one predictor (recall that it was 3%). This is in part due to the increase in the number of predictors with no commensurate increase in the number of cases.

The regression line is $Y' = -4.25 + 7.72(\text{GPA}) + 0.11(\text{cognitive ability})$. The printout shows that the *t* value for GPA is significant ($t = 4.01$, significance of 0.001), but for cognitive ability is not ($t = 0.92$, significance of 0.369). This means that GPA predicts starting salary above and beyond what cognitive ability does, but cognitive ability does not predict starting salary above and beyond GPA.

In addition, the beta values confirm that the predictive value of cognitive ability is questionable. That is, the relative strength of GPA (0.68) is quite high compared to that of cognitive ability (0.16). Thus, with this data set it would be concluded that the measure of cognitive ability does not enhance the prediction of starting salary, whereas GPA does.

Box 1.5 Multiple Regression Output from SPSS

Model Summary

R	*R-Square*	*Adjusted R-Square*	*Std. Error of Estimate*
0.72	0.52	0.46	4.3559

ANOVA

	Sums of Squares	*df*	*Mean Square*	*F*	*Sig.*
Regression	342.44	2	171.30	9.02	0.002
Residual	322.56	17	18.97		
Total	665.00	19			

Coefficients

Model	*Unstandardized b*	*Standard Error*	*Standardized (Beta)*	*t*	*Sig.*
Constant	–4.25	12.99		–0.33	0.747
GPA	7.72	1.93	0.68	4.01	0.001
Cognitive	0.11	0.12	0.16	0.92	0.369

There is a sample size problem here. Recall that there should be about 10 cases per variable in the equation. Because there are three variables, there should be about 30 cases but there are only 20 in this analysis. It is not that the regression program won't run—it will. It is up to the researchers to indicate that a lower than desirable sample size was used in the analysis and that, therefore, caution needs to be exercised so that the results are not overinterpreted. As scientists, it is convention to err on the side of being conservative in knowledge claims.

Correlation and regression will be used frequently in the coming chapters and thus a cursory review was deemed warranted at this point. If this brief overview was not sufficient, please see any number of introductory statistics textbooks to refresh more fully these topics.

Score Meaning. Raw scores on tests need to be interpreted. The numbers attached to raw scores are only meaningful in the context of a referent group of scores. For example, if I say, "I got a 15 on my history exam!" you don't know what that means—did I do well or poorly? This comparative information is called *normative* information. It is determined by knowledge about the referent group; in this case, you need to know how well the rest of the history class did on the exam to make my 15 meaningful.

Normative information is important for making sure that correct interpretations of scores are made. The larger and more representative the reference group to which a single score is compared, the more confidence can be placed in the interpretation of that score's meaning. For example, it would be better to compare my history exam mark of 15 with 1,000 students' marks over the last 10 years than to compare it with the marks of three classmates sitting around me.

This is why, for the more popular published tests, norm tables are provided. These tables have been created over many years by collecting large samples of data from many different test takers. These tables are sometimes broken down by demographic variables such as gender or age. This is so the test score interpreter can make a comparison of a score with the most appropriate demographic group. These normative samples are both very large and representative of the demographic characteristics of the group. An assumption in using norm tables is that the same test was used and administered under the same conditions as in the normative sample. Thus, it is up to test administrators to familiarize themselves with the administration protocol.

To make any raw score meaningful, it can be transformed into a distribution of meaningful, familiar values. The distribution most commonly known to social scientists is the standard normal distribution discussed earlier. It has a mean of 0.0 and a standard deviation of 1.0. This distribution is used because most individual differences are assumed to be normally distributed in the population. However, it is important to examine the degree to which this assumption is met in any sample data set. Luckily, most of the statistical procedures used in this text are robust (yield similar results to those found in normal distributions) to deviations of normality.

Converting a raw score (X) to a standard score (z score) based on the normal distribution is done via a simple transformation:

$$z \text{ score} = (X – \overline{X})/SD \text{ (standard deviation)}, \tag{1–11}$$

where the $\overline{X}$ and standard deviations are based on the sample.

Let's reconsider my history exam score of 15. If the mean of the class was 13 and the standard deviation was 2, then my raw score converts to a standard score of

$$z \text{ score} = (15–13)/2,$$

$$z \text{ score} = 1.0.$$

This means that if I look up my score of 1.0 in a distribution of normal scores, I see that I did better than about 84% of the class. Once I determine my z score, I can convert it to any other distribution where the mean and standard deviations are known. Some well-known distributions are the T score distribution with a mean of 50 and standard deviation of 10, or the Graduate Record Exam (GRE) distribution with a mean of 500 and standard deviation of 100. To make the conversions, simply use the following equation:

$$\text{New distribution score} = (z \text{ score} \times SD \text{ new}) + \text{Mean new}. \tag{1–12}$$

Assume someone has a *z* score of 0.43 on the GRE. The new GRE distribution score would be

$$\begin{aligned} \text{GRE} &= (0.43 \times 100) + 500, \\ &= (43) + 500, \\ &= 543. \end{aligned}$$

Assume someone has a *z* score of –1.5 on a test that will be converted to a *T* score. The new *T* distribution score would be

$$\begin{aligned} T\ \text{score} &= (-1.5 \times 10) + 50, \\ &= (-15) + 50, \\ &= 35. \end{aligned}$$

Figure 1.3 shows the normal distribution and the *T* score equivalents of some of the major values on the distributions. Thus, it can be seen that transforming scores simply means taking the values on one distribution and changing the values to reflect another distribution.

Another common conversion of raw scores is to percentile ranks. This is the percentage of individuals in the reference group earning a lower score than the score obtained. So, for example, if a score of 153 on a test is obtained and this is superior to 67% of the reference group, then the percentile rank is 67%. To calculate the percentile rank, the raw scores of the sample must be known as well as the number of individuals in the sample. Percentiles are not as commonly used as are standardized scores.

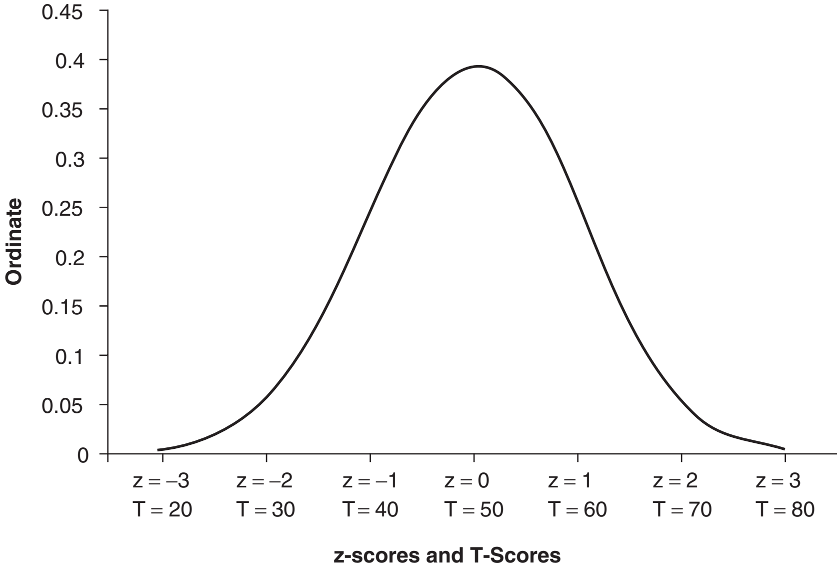

Figure 1.3 Normal Distribution Showing Selected *z* and *T* Scores

Quite a bit of time has been taken to review some of the information needed before embarking on the actual construction of a test. It is now time to make the transition to beginning the test development process. The first of these steps is to be clear on the construct to be measured. The rest of this chapter is devoted to that issue.

The First Step: Identifying the Construct

The first step in building any type of tool to assess individual differences is to identify the construct. The Webster's dictionary (Guralnik, 1976) defines a construct as "An idea or perception resulting from a synthesis of sense impressions." This is a useful definition, because it intimates that constructs are amorphous things; they are "ideas" and these ideas are a synthesis about a series of impressions. In other words, constructs are self-defined. The onus is on the test developer to convince the test user that the construct that is being measured is a reasonable assimilation and synthesis of ideas. Arguments are commonplace in the social sciences about "what we mean by construct *X*." One person's definition may not be the same as another's. If I ask an entire class to write down a definition of *success,* I will get as many different responses as there are students in the class. This means that it is unlikely that I will be able to create a test to assess success that will meet the expectations of all the students.

What is expected with a scale that is developed or used is that the individuals who respond to the items will provide information that will allow inferences to be made about the construct. Let's assume, for example, that we want to define the construct of being a team player in an organizational setting. Eventually, we will create a set of items that will, it's hoped, operationalize the construct of being a team player. For the moment, however, we'll concern ourselves with the issues of defining the construct itself.

It is helpful to have a list of what should be included and excluded from the construct. For example, in our assessment of being a team player, we'll restrict the construct to work settings; being a team player in sports, personal relationships, and so forth are not to be included in the domain of interest. Other aspects of what we want to include and what we want to exclude are also noted (see Table 1.5). This process makes explicit, for both the test developer and others as test consumers, what the measure will try to encompass and what it will not.

Links Between Constructs. For many tests, the goal is not just to test for the sake of testing. Instead, making inferences about the scores obtained on those tests is of interest. For example, the GRE is often used to make inferences about how well a test taker will do in graduate school. An assumption here is that the GRE assesses some cognitive skills that are needed to be successful in graduate school. The validity of this inference about GRE scores, then, is dependent on two things: the actual link between cognitive skills in graduate school and how well the GRE measures cognitive skills. In order to test this inference, actual numbers (data) must be

Table 1.5 Included and Excluded Aspects of Being a Team Player

Included	*Excluded*
Workplace examples	Sports examples
School projects	Personal relationships
Past experiences	Present circumstances
Outcomes	Personality conflicts
Progress to ends	Non-Western cultures
Effort expended	
Evaluation of results	

collected and will include GRE scores and some measure of success in graduate school. The strength of this link will be calculated using the actual data collected. Keep in mind, however, that the real purpose of the testing enterprise is to make inferences about the "true" link of cognitive skill and success in graduate school.

Figure 1.4 shows an example of what we may be likely to try to do with our scale of team performance. That is, we may want to predict an outcome, such as team effectiveness, and we want to use the amount of "team playerness" in teams as a predictor of that outcome. Recognize at the outset that a direct assessment of the true relationship between the amount of team playerness and team effectiveness is not possible (noted as the *hypothesis question* relationship in Figure 1.4). Instead, a *calculated value* between two measures of the constructs will be obtained. The measures of the constructs are ideally going to be accurate assessments of being a team player and team effectiveness. While the desired relationship is aimed for, calculated relationships inevitably fall quite short of that mark. An example will assist in making this clear.

Let's say we are in an organization that creates computer software. We want to know if being a good team player is related to team performance. There are an infinite number of ways to assess how much being a team player is part of the work group. One of the measures of the effectiveness of the team's work could be the number of errors that have to be debugged in the computer program. This is obviously only one of many potential ways to assess team effectiveness, but it will be helpful to keep things simple for the time being.

Now, information is collected about being a team player using a team player measure and about team effectiveness using a measure of the number of errors. We now have numbers to calculate the relationship between the two measured variables. It is apparent that the tighter the measurement linkages between the ideal constructs (shown in circles in Figure 1.4) and the measured variables (shown as squares in Figure 1.4), the better. *Better* in this context means more confidence in the inferences and knowledge claims about the link between being a team player and team effectiveness based on a calculated relationship between two measures.

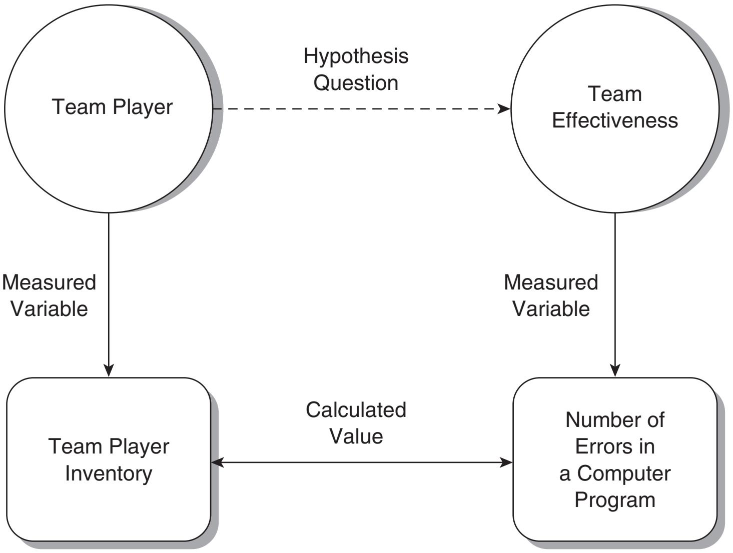

Figure 1.4 Linkages Between Hypothesized and Measured Relationships

Construct Cleanliness. Constructs are "clean" when they evaluate what they are supposed to; that is, the measurement links shown in Figure 1.4 are perfect insofar as the measured variables correspond 100% to the idea constructs. If a variable perfectly represents the construct, then measurement issues are not a concern. The trouble is that measured variables inevitably represent the idea construct imperfectly.

These imperfections come in two types: deficiency and contamination. A variable is deficient to the extent that the domain of interest is not covered. If I want to assess the extent that someone is likely to be a team player and I do not ask about that person's past experiences working on teams, my variable is likely to be deficient. If a group of 4th-graders is told that they will have a test on basic math skills but they are only given subtraction problems, the test is deficient in that addition, multiplication, and division problems have not been included.

Contamination of a construct by a measured variable is when the measure contains information that should not be part of the construct. If the team player assessment tool is administered to a team and they all fill it out together so that they all see each other's responses, "socially desirable" responses are more likely to occur rather than true responses. If a 4th-grade class is told that that they will be having a test on division but addition items are included, then the test is contaminated.

Contamination is easier to detect than is deficiency in any variable using various statistical procedures. Deficiency, however, has to be demonstrated rationally. If a construct seems to be missing something, finding that missing something usually comes from reviewing the existing theories and research or from practical knowledge about the construct.

Single Versus Multiple Constructs. An extremely important consideration in scale development and assessment is the extent to which the scale measures single or multiple constructs. This is not a simple matter. For example, volumes of writing and much work have gone into taking sides in the debate about whether or not intelligence is a multifaceted construct or a single construct. The evidence that both sides produce is logical and statistically sophisticated. The question, though, remains unanswered.

Some who have created scales to measure certain constructs have called a truce on this issue. For example, there are scales that measure facets of job satisfaction (e.g., satisfaction with pay, promotion, supervisor, etc.) and others that measure overall job satisfaction. Both are useful in different contexts. If an organization wants to assess if a new promotion system has had an effect on job satisfaction, then assessing "satisfaction with promotional opportunities" is more relevant than measuring other facets of job satisfaction (such as satisfaction with coworkers) or overall job satisfaction. On the other hand, if a new leadership team makes large structural changes in the organization, the members may be interested in the effects this might have on overall job satisfaction and therefore the overall measure would be more appropriate.

There is no right answer in the development of constructs as to whether the construct is unitary or multiple. It is better that a scale is developed with a clear idea first about whether one or multiple constructs are to be measured. Multiple constructs are more difficult to measure because, in addition to measuring them individually, how the constructs work together and relate to one another must be understood. This layer of complexity is best handled methodologically and statistically if it is posited to exist in advance of collecting any data.

Summary and Next Step

In this introductory chapter we have

a. reviewed why measurement is critical for science and why the problems associated with measurement in the social sciences pose unique problems,

b. provided a brief summary of some of the historical highlights of measurement,

c. reviewed the nomenclature around scales of measurement,

d. reviewed some of the basic premises of statistical analyses,

e. presented some of the common statistical procedures we'll be using in this text,

f. showed how scores are made meaningful by transforming them, and

g. presented the first step in developing any scale—defining the construct of interest.

The next step in the process of developing an instrument is to convert a construct into a series of stimuli (items) on which numerical information can be collected. This is the operationalization phase. It is time-consuming but, if done correctly, will save hours and days of time later on in the process.

Problems and Exercises

1. Recall a time when a test score had an impact on your life or on the life of someone you know. Describe what was measured, how it was measured, and how the score was used. Indicate the degree to which you felt that the test score was used appropriately and why.

2. An α level of 0.05 means what? What about an α level of 0.01, 0.001, 0.10, or 0.20?

3. Calculate the correlation coefficient by hand for the *X* and *Y* variables in Table 1.1.

4. Calculate the bivariate regression line of salary regressed on cognitive ability (data in Table 1.3). If you had a cognitive skill of 110, what would be your predicted starting salary level? Calculate the *R*-squared, adjusted *R*-squared, *F, Sb,* and *t* when regressing starting salary on cognitive ability. Interpret your findings.

5. At what level would the following variables be measured?
   a. Distances between towns
   b. Intelligence measured by an IQ test
   c. The rank ordering of members of a class based on height
   d. The numbering of those with blue eyes 1, brown eyes 2, green eyes 3, and other colors 4

6. If I obtain a score of 100 on a test that has a mean of 120 and a standard deviation of 10, what is my standard score? What would be my *T* score (mean of 50 and standard deviation of 10)?

7. Choose a construct that you are interested in finding more about. Here are a few examples to get you started thinking: civic-minded, athletic, studious, and humorous. Once you have selected your topic, create a chart like the one in Table 1.5. Share your ideas with your classmates.