Introducing correlational research

Native speakers of English who are learning Czech, Finnish, Spanish, or Polish are delighted to discover that the pronunciation of a word in these languages can be predicted with a high degree of accuracy by its spelling. Unfortunately, in English the correlation between spelling and pronunciation is not as close.

The American Heritage Book of English Usage: A practical and authoritative guide to contemporary English, 1996

The most merciful thing in the world … is the inability of the human mind to correlate all its contents.

H. P. LOVECRAFT, 1890–1937
American author of horror stories

In our daily lives, all of us think about how things fit together, how things are related, or in more technical terms, how they are correlated. Questions like the following come up every day:

1 Are big kids really faster runners?
2 Do blondes really have more fun?
3 Do students who get to school first in the morning get higher grades?
4 Do young males really drive faster than older males?

In each case, we could gather information in the form of numbers and examine the degree to which those numbers correlate. In the case of the four questions above, we could:

1 Study the correlation of students' heights to their times in the 100-meter dash (perhaps for a group of fifth and sixth graders).
2 Examine the degree to which blondeness (as rated by a panel of experts) correlates with the amount of fun the participants report having in life, as measured by a self-report questionnaire in which the respondents rate in a

variety of ways the amount of fun they have on average. (We could also do a before-and-after survey of people who dye their hair blonde.)

3 Investigate the relationship between the average number of minutes a certain group of high school students arrive early (or late) to school and compare their averages over a semester to their grade point averages (on a scale of 0 to 4).

4 Explore the relationship between the number of speeding tickets received by males in a particular year and the age of the drivers who got them.

Similarly, in formal research, one purpose is often to explore the degree to which two sets of numbers are related to each other in one way or another. Let's say you have developed a test you call the *New English Placement Examination* (NEPE) with the purpose of assigning students into the different levels of study in your ESL program. One interesting research question about the NEPE might be whether students' scores on the test are related to their scores on the *Michigan Test of English Language Proficiency* (MTELP). To investigate that question you might use the NEPE for placing students, then administer the MTELP one week later. To find out how much relationship there is between these two sets of scores, you could just line up the students' scores on the NEPE next to their MTELP scores as shown in Table 6.1.

Table 6.1 Example data comparing two sets of scores for a group of ESL students (www)

Name	*NEPE*	*MTELP*
Fahtima	100	98
Jose	94	92
Maria	86	89
Jaime	89	86
Abdullah	78	75
Noriko	76	79
Hans	64	69
Tatania	61	54
Jürgen	55	51

Clearly, just by examining the scores in Table 6.1, you can see that Fahtima and Jose have the highest scores on both tests and Tatania and Jürgen have the lowest. You can also see that the other students line up pretty much in the same order on both tests, but not in exactly the same order. In other words, you can see that the scores are related to some degree, but it is difficult to say exactly how much they are related. To get more precision, you will need to find out the CORRELATION, or degree of relationship, between the two sets of scores. This type of research is generally known as 'correlational research' (which we will define later).

Analyzing correlational data

How do we make sense of all those numbers? One way we can do so is to do CORRELATIONAL RESEARCH. This involves gathering and compiling data and then calculating a statistic called a CORRELATION COEFFICIENT. Correlation coefficients indicate the degree of relationship between two sets of numbers represented as the ratio of go-togetherness to total score variation. Thus correlation coefficients can range from 0.00 (if the ratio is zero, indicating absolutely no relationship) to 1.00 (if the ratio is perfect, indicating that there is a 100% relationship and that both sets of numbers are going in the same direction). For instance, the following numbers, representing, say the number of words spelled correctly in a spelling test of ten items, are perfectly correlated at 1.00 (because they are in a perfect relationship and going in the same direction):

Student	*Set A*	*Set B*
Marie	9	8
José	8	7
Jeanne	7	6
Hachiko	6	5
Raphael	5	4
Yuka	4	3
Hossein	3	2
Tamara	2	1
Hans	1	0

Naturally, all the numbers between 0.00 and 1.00 are also possible and, as they vary, they indicate differing degrees of relationship. So a correlation of .97 would be considered very high, and one of .09 would be very low, and others of .15, .27, .51, .67, .72, and .83 would be at different points in between.

Correlation coefficients can also range from 0.00 (again, for no relationship at all) to −1.00 (for a perfect relationship with the two sets of numbers changing together but in opposite directions). For instance, the following numbers, from, say, a different spelling test, would be perfectly correlated at −1.00:

Student	*Set A*	*Set B*
Marie	9	1
José	8	2
Jeanne	7	3
Hachiko	6	4
Raphael	5	5
Yuka	4	6
Hossein	3	7
Tamara	2	8
Hans	1	9

Naturally, all the coefficients between 0.00 and −1.00 are also possible and, as they vary, they indicate differing degrees of relationship. So a correlation of −.97 would be considered very high in a negative direction, and one of −.08 would be very low in a negative direction, and others of −.17, −.25, −.57, −.63, −.75, and −.84 would be at different points in between.

At this point, the important things to remember about correlation coefficients are that (a) the size of the number indicates the degree of relationship relative to a low of 0.00 and a high of 1.00 and (b) the sign (+ or −) indicates the direction of the relationship in a positive (i.e. same) direction or negative (i.e. opposite) direction.

Now that you understand how correlation coefficients work in general, let's look at how correlational research is done. Typically, three basic steps are followed in doing a correlational analysis:

(a) figure out what kind of scales you are dealing with,
(b) decide what kind of correlation coefficient to calculate, and
(c) calculate the appropriate correlation coefficient.

We will now consider each of these steps in more detail.