ORIGINAL EXAMPLE OF PERFORMANCE, ABILITY, AND MOTIVATION

Educational researchers have long argued that performance in school is a function of two factors: a student's motivation to perform well and his or her ability to perform well. This relationship is often expressed in the form of a multiplicative model, as follows:

$$\text{Performance} = \text{Ability} \times \text{Motivation} \tag{1}$$

The basic idea is that if a student lacks the cognitive skills and capacity to learn, then it does not matter how motivated he or she is; school performance will be poor. Similarly, a student can have very high levels of cognitive skills and the ability to learn, but if the motivation to work and attend to the tasks that school demands is low, then performance will be poor. The multiplicative relationship reflects this dynamic because, for example, if motivation is zero, then it does not matter what a person's score on ability is—his or her performance will always equal zero. Similarly, if ability has a score of zero, it does not matter what a person's motivation score is—his or her performance will always equal zero. Although this makes intuitive sense, the dynamics might be different from those implied by Equation 1, as we will now illustrate.

Our first step is to specify the metrics of the variables involved, since they do not have natural metrics. Performance in school might be indexed for individuals using the familiar grade-point average metric that ranges from 1.0 (all F's) to 4.0 (all A's), with decimals rounded to the nearest tenth (e.g., 2.1, 3.5). Ability might be indexed using a standard intelligence test that has a mean of 100 and a standard deviation of 15. Motivation might be indexed using a 10-item scale that asks students to agree or disagree with statements such as "I try hard in school" and "Doing my best in school is very important to me." A 5-point agree–disagree rating scale (1 = strongly disagree, 2 = moderately disagree, 3 = neither agree nor disagree, 4 = moderately agree, and 5 = strongly agree) provides the range of possible responses. The responses to each item are summed to yield an overall score from 10 to 50, with higher scores indicating higher levels of motivation.

Note that neither of these metrics takes on a value of zero. Hence, the dynamic of having "zero" ability or "zero" motivation discussed above cannot occur. Indeed, one might question whether there is such a thing as "zero" intelligence (i.e., a complete absence of intelligence). Is a psychological zero point on this dimension even possible? Suppose we decide that although a complete absence of intelligence is not theoretically plausible, a complete absence of motivation to do well in school is plausible. One way of creating a motivation metric with a zero point is to subtract a score of 10 from the original motivation metric. Before this operation, the motivation metric ranged from 10 to 50. By subtracting 10 from the metric, it now ranges from 0 to 40, which includes a zero point.

However, there is a problem with this strategy. Just because we can mathematically create a zero score on the motivation scale by subtracting 10 from it, this does not mean that the score of zero on the transformed scale reflects a complete absence of motivation on the underlying dimension of motivation. What evidence do we have that this is indeed the case? Perhaps a score of zero on the new metric actually reflects a somewhat low level of motivation but not a complete absence of it. The issue of mapping scores on a metric onto their location on the underlying dimension they represent is complex, and consideration of how to accomplish this is beyond the scope of this book. We will work with the original metric of 10–50 and not make explicit assumptions about where on the underlying motivation dimension these scores locate individuals. We suspect that, based on the content of the items, students who score near 50 are very highly motivated to perform well, and students who score near 10 are very low in (but not completely devoid of) motivation to perform well. But a separate research program is required to establish such assertions (Blanton & Jaccard, 2006a).

Suppose that a student has a score of 100 on the IQ test and a score of 30 on the motivation test. Using Equation 1, multiplying the ability score by the motivation score, we obtain 100 × 30 = 3,000, and we would predict a GPA of 3,000! Of course, this is impossible because a student's GPA can range only from 1.0 to 4.0. We need to introduce one or more adjustable constants to Equation 1 to accommodate the metric differences and to make it so that a predicted GPA score falls within the 1.0–4.0 range. For example, if we let P stand for performance, A for ability, and M for motivation, then we can allow for the subtraction of a constant from the product to make an adjustment in metric differences, modifying Equation 1 as follows

$$P = (A)(M) + a$$

where a is an adjustable constant whose value is estimated from data. Note, for example, if $a = -2{,}997$, then this is the same as subtracting 2,997 from the product of A and M. But perhaps subtracting a constant is not enough to account for the metric differences. For example, a score of 120 on the IQ test coupled with a score of 50 on the motivation test would yield a product value of 6,000, and subtracting a value of 2,997 from it would still produce a nonsensical GPA. A second scalar adjustment we might use is to multiply the product term by a fractional adjustable constant, which yields the general equation

$$P = b(A)(M) + a$$

where b is a second adjustable constant (in this case, b would be a fraction) designed to deal further with the metric differences. Its value also is estimated from data. The terms on the right-hand side of this equation can be rearranged to yield

$$P = a + b(A)(M) \qquad (2)$$

Equation 2 is simply a linear function, so performance is assumed to be a linear function of the product of (*A*)(*M*). Not only do the constants *a* and *b* take into account the different metrics, but the value of *b* also provides substantive information as well; namely, it indicates how much change in performance (GPA) one expects given a 1-unit increase in the value of the product term (*A*)(*M*).

Figure 1 plots the relationship between performance and motivation at three different levels of ability based on Equation 2, where values of *a* and *b* have been empirically determined from data collected for a sample of 90 students. In this example, *a* = –2.0 and *b* = .0015. The slope of *P* on *M* for any given value of *A* is *bA*. There are several features of this plot worth noting. First, note that the effect of motivation on performance is more pronounced as ability increases. This is evident in the steeper slope (*bA* = .165) for the two variables when the ability score is 110 as compared with the slope when the ability score is 100 (*bA* = .150), and, in turn, as compared to the slope when the ability score is 90 (*bA* = .135). These differences in slope may seem small but they are probably substantial. For example, when the ability score is 110, a 10-unit change in motivation is predicted to yield a (.165)(10) = 1.65-unit change in GPA; when the ability score is 100, a 10-unit change in motivation is predicted to yield a (.150)(10) = 1.50-unit change in GPA; when the ability score is 90, a 10-unit change in motivation is predicted to yield a (.135)(10) = 1.35-unit change in GPA.

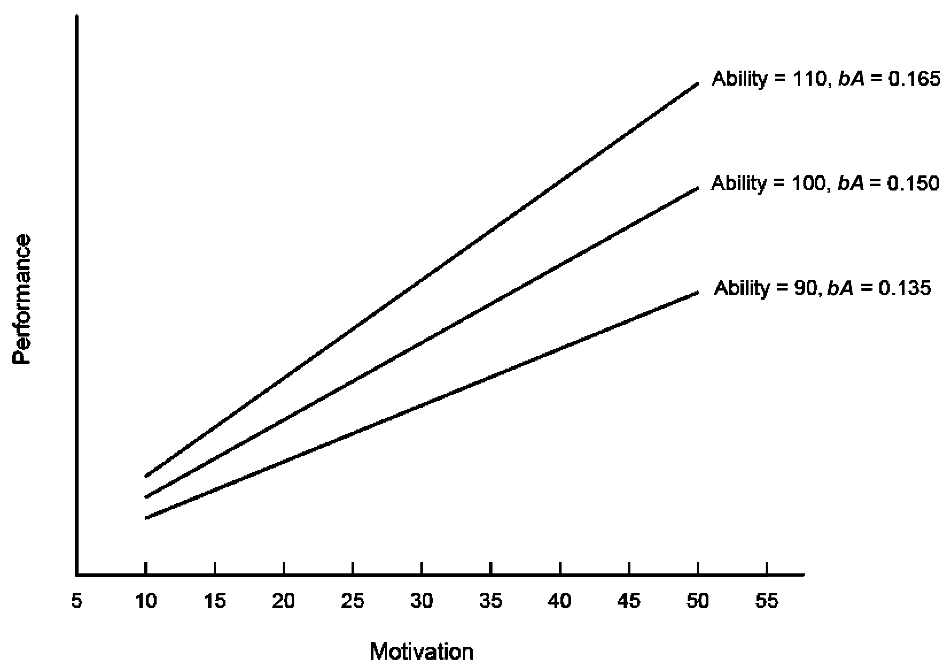

FIGURE 1. Example for Performance, Ability, and Motivation

Second, note that at each of the different levels of ability (90, 100, and 110), the relationship between motivation and performance is assumed to be linear. Is this a reasonable assumption? Perhaps not. Perhaps the relationship between performance and motivation at a given ability level is better captured by an exponential function. For example, when motivation is on the low end of the motivation metric, increasing it somewhat may not have much impact on performance—it will still be too low to make a difference on performance. But at higher levels of the motivation metric, increasing it will have an impact on performance. Or perhaps a power function is applicable. Power functions have the same dynamic as the exponential function, but they "grow" a bit more slowly. Or perhaps an S-shaped function applies, with floor and ceiling effects on performance occurring at the low and high ends of motivation, respectively.

The multiplicative model specified by Equation 1 assumes what is called a *bilinear interaction* between the predictor variables; that is, it assumes that the relationship between the outcome and one of the predictors (in this case, motivation) is always linear no matter what the value is of the other predictor (in this case, ability). To be sure, the value of the slope for the linear relationship between P and M differs depending on the value of A (as noted earlier), but the function form is assumed to be linear. One can modify the model to allow for a nonlinear relationship between performance and motivation at different levels of ability, say, in accord with a power function, as follows

$$P = a + b(A)(M^c) \tag{3}$$

where c is an adjustable constant whose value is estimated from data.

Another notable feature of Figure 1 is that at the lowest value of motivation, there is a small degree of separation between the three different lines. The amount of separation between the lines reflects the differences in the effect of ability (at values of 90 vs. 100 vs. 110) on performance when motivation is held constant at the same value. But perhaps the amount of separation should be a bit more or a bit less than what is modeled in Figure 1. The equation can be further modified to allow for a different amount of separation between the lines than what Equation 3 implies, as follows:

$$P = a + b(A)(M^c) + dA \tag{4}$$

where d is an adjustable constant whose value is estimated by data. The logic of adding this term is developed in the Appendix and is not central to our discussion here. The main points we want to emphasize are the following:

1. The theoretical representation in Equation 1 has nontrivial conceptual implications because it takes the strong stand that the relationship between performance and the

predictor variables is captured by the dynamics of a bilinear interaction. In fact, the interaction may have a different functional form.

2. When building a mathematical model, the metrics of the variables usually have to be taken into account (although our next example illustrates a case where this is not necessary).

3. There may be multiple features of the model (e.g., the separation between curves at different levels of the component terms as well as the shape of these curves) that must be specified that are not always apparent in simple representations such as Equation 1.

The fact is that the often presented model of Performance = Ability × Motivation is under-developed, and applying principles of mathematical modeling helps to produce a better-specified theory that makes implicit assumptions explicit and highlights complexities that should be taken into account. The Appendix of this document develops modeling strategies for this example in more detail and illustrates a substitution principle for building mathematical models.

AN EXAMPLE USING COGNITIVE ALGEBRA

Another example of using mathematical models to represent social phenomena involves models of cognitive algebra. This example illustrates how the implications of a mathematical representation can be pursued without recourse to such things as adjustable constants and complex modeling of data.

Suppose we describe the personal qualities of a political candidate to a person that he or she has not heard of by providing the person with three pieces of information. Suppose that the three pieces of information are all quite positive (e.g., the candidate is said to be honest, smart, and empathic). For purposes of developing this example, suppose we can characterize how positive each piece of information is considered to be using a metric that ranges from 0 to 10, with higher numbers reflecting higher degrees of positivity. We refer to the positivity of a piece of information as P_k, where k indicates the specific piece of information to which we are referring: P_1 refers to the perceived positivity of the first piece of information, P_2 refers to the perceived positivity of the second piece of information, and P_3 refers to the perceived positivity of the third piece of information. Suppose we want to predict how favorable a person will feel toward the candidate based on these three pieces of information. If we let F refer to a person's overall feeling of favorability toward the candidate, with higher values indicating higher levels of favorability, then one model that describes the impact of the information is the following:

$$F = P_1 + P_2 + P_3 \quad (5)$$

This model is a simple summative function that specifies that the overall feeling of favorability toward the candidate is the sum of the judged positivity of each individual piece of information (we ignore, for the moment, the metric of *F* and the issue of adjusting for metric differences). Equation 5 can be stated in more general form using summation notation as follows:

$$F = \sum_{i=1}^{k} P_i$$

where *k* is the number of pieces of information, in this case 3.

Now suppose that instead of a summative function, an averaging function is operating. That is, the overall feeling of favorability is the *average* of the positivity of the information presented rather than the sum of it. In this case, Equation 5 becomes

$$F = (P_1 + P_2 + P_3)/3 \qquad (6)$$

and this can be represented more generally in summation notation as

$$F = \left(\sum_{i=1}^{k} P_i \right) \Big/ k$$

What are the implications of specifying the function as being summative versus averaging in form? It turns out, they are considerable. Let's explore the summation model first. Suppose a person judges the positivity values of the three pieces of information as 8, 8, and 8, respectively. The overall feeling of favorability toward the candidate will be 8 + 8 + 8 = 24. Now suppose we describe a second candidate to this person using the same three pieces of information but we add a fourth descriptor to them (cunning), that is judged to have a positivity value of 4. According to the summation model, the overall feeling of favorability toward this new candidate will be 8 + 8 + 8 + 4 = 28, and the person will prefer the second candidate to the first candidate. Psychologically, it is as if the second candidate brings all the same qualities as the first candidate (i.e., P_1, P_2, and P_3) and then "as a bonus," you get a fourth positive attribute as well (P_4). Hence, the person prefers the second candidate to the first candidate.

Now consider instead the averaging function. The overall feeling toward the first candidate is predicted to be (8 + 8 + 8)/3 = 8.0 and the overall feeling toward the second candidate is said to be (8 + 8 + 8 + 4)/4 = 7.0. In the averaging model, exactly the reverse prediction is made in terms of candidate preference; namely, the person now will prefer the first candidate to the second candidate. Psychologically, the first candidate has nothing but very positive qualities, whereas the second candidate has very positive qualities but also some qualities that are only somewhat positive. The person prefers the first candidate, who has nothing but very positive qualities, to the second candidate, who has very positive qualities but also moderately positive qualities.

Which function better accounts for the impressions people form? It turns out that this can be evaluated in a simple experiment in which two candidates would be described, one with three very positive qualities (Candidate A) and a second with three very positive qualities and a fourth moderately positive quality (Candidate B). Participants would then be asked to indicate which of the two candidates they prefer. The summation model predicts that participants should prefer Candidate B to Candidate A, whereas the averaging model predicts that participants should prefer Candidate A to Candidate B. One can differentiate the two models empirically by conducting the above experiment and determining which candidate tends to be preferred. This is a simple experiment without complex modeling. If the results showed that people tend to prefer Candidate A to Candidate B, then this would be consistent with (but not proof of) a summative process rather than an averaging process. If the results showed that people tended to prefer Candidate B to Candidate A, then this would be consistent with (but not proof of) an averaging process rather than a summative process. Which process operates has implications for the design of political campaigns and advertising strategies to sell products. For example, if an advertising campaign adds to a person's cognitions a moderately positive piece of information about a product that is already quite positively evaluated, in the case of the averaging model, the advertisement should backfire and lower evaluations of the target product, thereby adversely affecting sales.

The literature on impression formation has extended these simple models of "cognitive algebra" to more complex model forms. For example, it is almost certainly the case that some information is more important to people in forming impressions than other information. As such, it makes sense to weight each piece of information by its importance to the individual. Equation 5 can be modified to include such weights, as follows:

$$F = w_1P_1 + w_2P_2 + w_3P_3 \quad (7)$$

where w_i is the importance of information i to the individual. Note that Equation 5 is a special case of Equation 7, namely the case where $w_1 = w_2 = w_3 = 1$. Expressed in summation notation, Equation 7 can be represented as

$$F = \sum_{i=1}^{k} w_i P_i$$

For the averaging model, introducing importance weights yields the following:

$$F = (w_1P_1 + w_2P_2 + w_3P_3)/(w_1 + w_2 + w_3) \quad (9)$$

Equation 9 can be restated using summation notation as

$$F = \sum_{i=1}^{k} w_i P_i \Big/ \sum_{i=1}^{k} w_i$$

By extending the logic of algebraic models to the domain of “cognitive algebra” (which uses the premise that mental operations can be modeled by simple algebra), a great many insights into human information processing have been gained. Some of this research has involved simple experiments that pit competing predictions of different algebraic models against one another, whereas other research has taken the path of more complex math modeling with adjustable constants, error terms, and the like.

Parenthetically, the research literature finds support for both the summation and averaging models. In some contexts, people average the implications of information, whereas in other contexts, they sum it. There also are individual differences in these tendencies, with some people tending to average information in general whereas others tend to sum it in general. There are contexts for which simple summation or averaging models do not hold, and more complex combinatorial models are required to capture the integration dynamics.

APPENDIX: MODELING ISSUES FOR PERFORMANCE EXAMPLE

This appendix describes details for the example modeling the effects of ability and motivation on performance, where the relationship between performance and motivation is nonlinear instead of linear at a given level of ability. We assume the reader is versed in standard statistical methods and psychometric theory. We illustrate the case first where motivation is assumed to impact performance in accord with a power function, with the shape of the power function changing as a function of ability. Then we mention the case where the relationship between performance and motivation is assumed to be S-shaped, with the form of the S varying as a function of ability.

We build the power function model by first positing that performance is a power function of motivation,

$$P = a + bM^{c} \tag{A.1}$$

where a and b are adjustable constants to accommodate metrics and c is an adjustable constant to isolate the relevant power curve in light of a and b. According to the broader theory, the effect of motivation on performance varies depending on ability (e.g., when ability is low, increases in motivation will have negligible effects on performance, but when ability is moderate to high, increases in motivation will have a more substantial impact on performance). Stated another way, the shape of the power curve will differ depending on the level of ability of students, such that the value of c is some function of A. In addition, it is likely the case that the adjustable constants a and b vary as a function of A. To simplify matters and to develop the underlying logic, we will assume that c is a linear function of A, that a is a linear function of A, and that b is a linear function of A. This yields the equations

$$c = d + fA$$
$$a = g + hA$$
$$b = i + jA$$

where c, d, f, g, h, i, and j are adjustable constants that conform to the respective linear models. Using substitution principles, we can substitute the right-hand side of these equations into A.1, which yields

$$P = (g + hA) + (i + jA)(M)^{(d + fA)}$$

Expanding, we obtain

$$P = (g + hA) + iM^{(d + fA)} + jAM^{(d + fA)}$$

We can rewrite this equation using the more familiar symbols of *a* and *b* for adjustable constants in regression analysis:

$$P = a + b_1A + b_2M^{(b_3 + b_4A)} + b_5AM^{(b_3 + b_4A)}$$

This model can be fit to data and the values of the adjustable constants estimated using nonlinear regression algorithms in SPSS or some other statistical package. The adjustable constants are amenable to interpretation, but we forgo explication of this here. Additional interpretative complications present themselves if the metrics involved are arbitrary, but we do not pursue such matters here either.

One intuitive way of seeing the implications of the function once the values of the adjustable constants are estimated is to calculate predicted scores that vary M by 1 unit at select values of A. These can be graphed and then subjected to interpretation.

An alternative approach to modeling the data that uses methods that are more familiar to social scientists is to use polynomial regression. In this approach, performance is assumed to be a quadratic function of motivation. Although the full quadratic curve most certainly is not applicable (because it is U-shaped), the part of the curve that forms the right half of the "U" could apply. The model includes adjustable constants to isolate this portion. We begin by writing a model where performance is a quadratic function of motivation

$$P = a_1 + b_1M + b_2M^2 \tag{A.2}$$

and the adjustable constants in this equation (the intercept and the regression coefficients) are modeled as being a linear function of ability (we could use a nonlinear function, but for the sake of pedagogy, we assume a linear function), yielding

$$a_1 = a_2 + b_3A$$
$$b_1 = a_3 + b_4A$$
$$b_2 = a_4 + b_5A$$

Using the substitution principle, we substitute the right-hand sides of these equations for their respective terms in Equation A.2, which produces

$$P = (a_2 + b_3A) + (a_3 + b_4A)M + (a_4 + b_5A)M^2$$

Expanding this yields

$$P = a_2 + b_3A + a_3M + b_4AM + a_4M^2 + b_5AM^2$$

Rearranging and relabeling the constants to conform to more traditional notation yields the model

$$P = a + b_1A + b_2M + b_3AM + b_4M^2 + b_5AM^2$$

This model can be fit using standard least squares regression.

To model an S-shaped function, one can stay with polynomial regression but extend the logic to a cubic function. The basic idea is to express performance as a cubic function of motivation

$$P = a_1 + b_1M + b_2M^2 + b_3M^3$$

and then to model the adjustable constants as a function of *A*. Finally, use the substitution method to derive the more complex generating function.

Alternatively, one can use a logistic function to capture the S shape and then model the adjustable constants within it as a function of *A*. This approach requires the use of nonlinear algorithms in estimating the adjustable constants.