# 4

# USING THE SIMPLIFIED RETICULAR ACTION MODEL NOTATION

It is fairly well-known that a specific matrix algebra formulation such as structural equation modeling (SEM)–linear structural relations model (LISREL; of Exhibit 3.2) is not the only way to create model expectations (McArdle, 2005). Without any loss of enthusiasm for the basic premises of SEM, it is fair to say that there were several knowledgeable researchers who suggested that the LISREL concept and computer program was not actually the best way to deal with such problems, and Roderick P. McDonald (1980) was among the most vocal (Loehlin, 1987, 1998, 2004; McDonald, 1985b). In McDonald's alternative approach, the newly available computer programming he created allowed an unrestricted level of higher order common factors that he termed *covariance structure analysis* (COSAN). The use of the resulting concepts and computer program produced exactly the same numerical values as LISREL, but it was not nearly as popular as LISREL, and this was unfortunate because COSAN was free software. The notation we will use here—the

http://dx.doi.org/10.1037/14440-005
*Longitudinal Data Analysis Using Structural Equation Models*, by J. J. McArdle and J. R. Nesselroade

reticular action model (RAM)—is slightly different, because it uses a simplified algebra to match the diagrams (see Figure 2.4).

As stated earlier, it is not necessary to use any particular algebraic notation as long as the general SEM principles are followed. Nevertheless, our own interest in path diagrams also led to another innovation in the calculation schemes. Our prior use of COSAN proved instrumental in asserting that the eight-matrix notation, and the entire LISREL concept, was not the only correct way to carry out SEM (Horn & McArdle, 1980). In fact, a practical comparison of COSAN and LISREL applications led directly to the simplification known as RAM theory (McArdle, 1979, 2005; McArdle & McDonald, 1984).

## RAM ALGEBRAIC NOTATION

It became clear to us (McArdle, 1978; McArdle & McDonald, 1984) that only three model matrices were necessary to consider any structural equation model: (1) a filter matrix (**F**) of completely fixed ones and zeros designed to distinguish the manifest variables from the latent variables in a model, (2) an arrow matrix (**A**) of potential one-headed arrows (regression coefficients and means) based on directional hypotheses, and (3) a sling matrix ($\mathbf{\Omega}$) of potential two-headed arrows based on nondirectional hypotheses. This resulting set of parameters is the RAM notation (McArdle & McDonald, 1984) with vectors (**p**, **u**, **q**) and model matrices (**A**, $\mathbf{\Omega}$, **F**). This is more simply termed *RAM specification* (McArdle, 2005), and we will use here it only because it matches the path diagrams exactly. RAM is fully defined in Exhibit 4.1, and the resulting expectations are given in Exhibit 4.2.

## RAM GRAPHICS

Perhaps it is now obvious that this path diagram (e.g., Figure 2.1 or 2.4) are not plots of the raw data but topographical representations of some of the assumptions in the model of analysis; that is, the residual has no mean, no correlation with the predictor, and so on. These path diagrams do not say anything about the distribution requirements for the predictor X, primarily because X can have any distribution. But it also does not say anything about the distribution of the residuals, although we do know that the unobserved residual scores (*e*) need to be "normally distributed" for the statistical tests to be exact. Nevertheless, one might reasonably ask, do we need to know all this just to define a simple linear regression? The complete answer to this question really depends on what we are going to do next.

EXHIBIT 4.1
A Generalized Set of Rules Used To Define a Model
for Any Structural Equation Model

Distinguish types of variables: There are only observed variables ($M$, drawn as squares) and unobserved variables ($L$, drawn as circles). The summation of all variables is a list (of size $V = M + L$).
Distinguish type of parameters: There are only directed relationships (**A**, drawn as one-headed arrows) and undirected relationships ($\mathbf{\Omega}$, drawn as two-headed slings).
Define the filter matrix **F**: Create a ($M$ by $V$) matrix of zeros, and enter unit values where the variables that is observed (in the row) is give the same name as the variable in the model (in the column).
Define the arrow matrix **A**: Create a square but nonsymmetric ($V$ by $V$) matrix of zeros, and enter unit values where any variable is the outcome (in the row) of any other variable in the model (in the column).
Define the sling matrix $\mathbf{\Omega}$: Create a symmetric ($V$ by $V$) matrix of zeros, and enter unit values where the variable in the model is connected (in the row) to any other variable in the model (in the column).

*Note.* From *RAMpath: A Computer Program for Automatic Path Diagrams* (pp. P15–P18), by J. J. McArdle and S. M. Boker, 1990, Hillsdale, NJ: Lawrence Erlbaum Publishers. 

One caveat is that, because we have introduced the constant (triangle labeled 1) as part of this expectation, we end up with expectations for a combination of both a mean vector and a covariance matrix, or a matrix of first and second moments. This kind of calculation is actually well-known by another name; we have created an average sums of squares and cross-products (SSCP) matrix expectation (Bock, 1975), and this is sometimes referred to as a *matrix of moments*. But we will use this matrix as a group summary in our basic fitting algorithm. However, since we will often be thinking about means and covariances separately to some degree, we need a convenient way to separate out the expected mean vector from this one. This can be accomplished in many ways. This is obviously the mean square plus the

EXHIBIT 4.2
Specifying Model Expectations for Any SEM
Using RAMpath Notation

Calculate all model effects **E**: Calculate the square matrix of total effects from one variable to another by writing $\mathbf{E} = (\mathbf{I} - \mathbf{A})^{-1} = \mathbf{I} + \mathbf{A} + \mathbf{A}^2 + \mathbf{A}^3 \cdots + \mathbf{A}^k$
Calculate all model expectations: Calculate the symmetric (V by V) matrix of all model expectations as $\mathbf{SSCP(V)} = \mathbf{E}\,\mathbf{\Omega}\,\mathbf{E}'$
Calculate observed model expectations: Calculate the symmetric (M by M) matrix of all model expectations as $\mathbf{SSCP(M)} = \mathbf{F}\,\mathbf{SSCP(V)}'\,\mathbf{F}'$ or $\mathbf{F}\,\mathbf{E}\,\mathbf{\Omega}\,\mathbf{E}'\,\mathbf{F}'$

covariances (i.e., the SSCP) listed in the prior equations. This SSCP matrix is used to summarize the expected information about the model-based raw scores. In this format it seems easier to see that the use of a constant has zero effect on the expected covariances, and the two-headed unit value on the constant is included to preserve the needed cross-product.

Given that we use RAM notation and we are sincere about completing our diagrams, we do not need any further steps. RAM notation represents the necessary and sufficient conditions for any structural equation model (for proof, see McArdle & McDonald, 1984). Once again, we should not forget that the main point of RAM notation is that these three matrices also were based on a one-to-one identity (or isomorphism) with the path analysis graphics (for details, see McArdle, 2005; McArdle & Boker, 1990). In various demonstrations, we have showed how this three-matrix approach produced exactly the same values as the eight-or-more-matrix approach (e.g., McArdle, 2005; McArdle & McDonald, 1984). This approach, in fact leads to a general statement of the model reconstruction (Exhibit 4.2) with an associated set of tracing rules (in Exhibit 4.3). We do not give details about these issues because these are simply reflections of the algebra.

### EXHIBIT 4.3
### A Generalized Set of Tracing Rules for Any Structural Equation Model

The following are after Wright (1918, 1921)—assuming RAMpath diagrams—by McArdle & Boker, 1990:

The RAMgraph: When defined using RAMpath rules can be briefly described in terms of a set of actions that represent the nonlinear structure of the model expectations.

The effect matrix **E** contains actions that are one-way tracings in a RAMgraph. Each element (**e**{x . . . w}) is the action from any variable v(w) to any variable v(x) that passes through a one-way asymmetric sequence of r > 0 adjacent variables connected by r – 1 one-headed arrows **a**{i,j}.

Bridges **B** are sequentially defined two-way movements or tracings in the RAMgraph. The element **b**{x . . . w:z . . . y} is a tracing from any initial variable c(x) to any terminating variable v(y) that passes through a symmetric sequence of (a) one backward action **e**{x . . . w} tracing backward from v(x) to v(w) connecting r > 0 variables, (b) one and only one sling (**s**{w,z}) connecting two mediating variables v(w) and v(z), and (c) one forward action (**e**{y . . . z} = **e**{z . . . y}′) traced from V(z) to v(y) connecting q > 0 variables.

Connections **C**: The complete linkage between pairs of variables is the sum of all nonparallel bridges between two variables. Thus, **c**{x,y} = **b**{x . . . w: z . . . y} **s**{y,z} **b**{z . . . y: w . . . x} also is termed the average cross-product or sums-of-squares. The bridges are nonparallel if they may be stated in a different notational sequence even though they may contain the same elements (i.e., **b**{x . . . w: z . . . y} NE **b**{x . . . z: w . . . y}.

*Note.* From *RAMpath: A Computer Program for Automatic Path Diagrams* (p. P11), by J. J. McArdle and S. M. Boker, 1990, Hillsdale, NJ: Lawrence Erlbaum Publishers. Copyright 1990 by Taylor & Francis. Adapted with permission.

We are suggesting that the existing SEM computer programs could be quite useful (e.g., LISREL, COSAN; see Exhibit 3.1), but all available matrices in these programs were not needed because only these three-parameter matrices were needed to produce all the correct model expectations ($\mathbf{\Sigma}$). The big benefit of this simplified programming is that it made exceedingly complex models relatively easy to consider (e.g., Grimm & McArdle, 2005; McArdle, 2005; McArdle & Hamagami, 2003). Nevertheless, any structural equation model only consists of variables that are either measured (squares) or not (circles), and relationships that are either directed (arrows) or not (slings). All other statements about the reasons why a specific approach should be used (i.e., combining econometrics with psychometrics) could still be useful, but they are certainly not essential. Of most importance, because this separation of measurement and structural model does not always lead to the best longitudinal analyses—as with the earlier COSAN concept, it was no surprise that the RAM concept was not uniformly recognized by the scholars who had put so much time and energy into this incredibly complex LISREL concepts, notation, and programming. Perhaps the original RAM names were unfortunate choices, but the same could be said for the LISREL concepts (McArdle, 2005). Nothing was necessary but it was useful.

## USING PATH DIAGRAMS AND RAM ALGEBRA

One very useful feature of SEM—the path diagram—was originally thought to be only a conceptual device. But the close relation of the algebraic equations to the subsequent path diagram is important, because it is useful, and may explain some of the popularity of SEM among novices (McArdle, 2005). However, because the path diagram does seem to convey information we, like others before us, will use it rather extensively here (O. D. Duncan, 1975; Heise, 1974). However, and unlike the others, in any path diagram used here, observed variables are drawn as squares, unobserved variables are drawn as circles, and (1) a constant (needed when means or intercepts are used) is included as a triangle. Using this notation, a path diagram of the traditional model of simple linear regression is depicted in Figure 2.4. This model can now be seen to have three additional variables: (2) an observed outcome (Y), (3) an observed predictor (X), and (4) an unobserved residual (*e*). We emphasize that there are actually four variables in this model, but the constant is usually not counted. The model also has three basic parameters: (1) a slope ($\beta$) indicating the difference in the expected outcome for every one-unit difference in the predictor, (2) the variance of the predictor (labeled $\sigma_e^2$), and (3) the variance of the residual (labeled $\psi^2$). From these last two variances we can calculated the explained variance in Y due to X (i.e., $R^2 = [\sigma_e^2 - \psi^2])/\sigma_e^2$), but the explained variance is not strictly a parameter in the basic model (we do not use it in computer scripts; see the companion book).

When we notice that the path diagram has only a few elements, it seems natural to do the same with the algebra (see Exhibit 4.2; McArdle, 2005; McArdle & McDonald, 1984). As a caution, what follows next is a brief algebraic presentation of RAM notation, and it is clearly not needed by everyone. In fact, we will really only use this once in the whole book, and we will concentrate our efforts on the path diagrams instead. However, the algebra and the diagrams are completely consistent, and the algebra really did come first, so all is not lost by restating the obvious features of the simple regression model, and this is done in Exhibits 4.4 and 4.5.

These kinds of path diagrams can be conceptually useful devices for understanding basic concepts or tools, and we use them often here. They also are conceptually useful because they allow a potentially complex multifaceted theory to be portrayed in a single display. These diagrams are also

EXHIBIT 4.4
Specifying a Regression Using RAMpath Notation

Assuming we would like to create a simplified version of the LISREL model matrices so we can write any model that is properly drawn (i.e., using the RAMpath rules of path analysis). To do so, we first simply define all variables in the model (**p**) as a list or vector:

(1) $$\mathbf{p} = [\mathrm{Y}\ \mathrm{X}\ 1\,e],$$

and we can then write the linear relationships among all variables in a general way as the model as

(2) $$\mathbf{p} = \mathbf{p}\,\mathbf{A} + \mathbf{u}$$

where **u** is introduced as a set of unknown residuals, and we define a square and nonsymmetric matrix of regressions or one-headed arrows (**A**) in this simple regression example as

(3) $$\mathbf{A} = \begin{pmatrix} |0 & \beta_1 & \beta_0 & 1| \\ |0 & 0 & 0 & 0| \\ |0 & 0 & 0 & 0| \\ |0 & 0 & 0 & 0| \end{pmatrix}$$

and we can verify that there are exactly four non-zero one-headed arrows in the diagram, but their placement in this matrix (as column variables into row variables) is defined by the order they were placed in the **p** vector (in Equation 1).

EXHIBIT 4.5
Specifying Regression Expectations Using RAMpath Notation

At this point it is generally useful to restate the relationships among all variables (Exhibit 4.4, Equation 2) as

(1)
$$\begin{aligned} \mathbf{p} &= \mathbf{p}\,\mathbf{A} + \mathbf{u}, \\ \mathbf{p} - \mathbf{p}\,\mathbf{A} &= \mathbf{u}, \\ \mathbf{p}(\mathbf{I} - \mathbf{A}) &= \mathbf{u}, \\ \mathbf{p} &= (\mathbf{I} - \mathbf{A})^{-1}\mathbf{u}, \\ \mathbf{p} &= \mathbf{E}\,\mathbf{u}, \end{aligned}$$

where we introduced a square identity matrix (**I**) and the inverse operator (–1) to contain a matrix of total effects (**E**), which for this simple example is

(2)
$$\mathbf{E} = (\mathbf{I} - \mathbf{A})^{-1} = \begin{pmatrix} |(1-0) & \beta_1 & \beta_0 & 1 & |^{-1} = & |1 & \beta_1 & \beta_0 & 1| \\ |0 & (1-0) & 0 & 0 & | & |0 & 1 & 0 & 0| \\ |0 & 0 & (1-0) & 0 & | & |0 & 0 & 1 & 0| \\ |0 & 0 & 0 & (1-0) & | & |0 & 0 & 0 & 1| \end{pmatrix}$$

While this can be an awfully confusing step (and is obviously trivial here because $\mathbf{E} = \mathbf{I} + \mathbf{A}$), what we have generally done is to rewrite all model variables (in **p**) as a linear function of their unknowns (in **u**). This now allows us to define the expected values of the unknowns ($E\{\mathbf{u}\,\mathbf{u}'\}$) using the placement of the two-headed slings ($\Omega$) defined here as a lower triangular matrix with non-zero elements. In the simple regression example this is

(3)
$$E\{\mathbf{u}\,\mathbf{u}'\} = \Omega = \begin{pmatrix} |0 & & & \text{sym.}| \\ |0 & \sigma_x^2 & & | \\ |0 & 0 & 1 & | \\ |0 & 0 & 1 & \psi_e^2| \end{pmatrix}$$

At this point we can verify that there are exactly three non-zero two-headed arrows in the diagram, and their placement in this matrix is defined by the order that we placed them in the **p** vector.

To complete the typical model specification, it is useful to separate (or filter out) the observed variables (squares or triangles) from the unobserved variables (circles). This is especially useful because we want to compare our model expectations (see below) with all *observed variables.* To do so, we first write the as subset of all **p** as

(4a)
$$\mathbf{q} = \mathbf{F}\,\mathbf{p}.$$

(*continues*)

EXHIBIT 4.5
Specifying Regression Expectations Using RAMpath Notation *(Continued)*

To get from the observed **q** from all **p**, we can write a *filter matrix* (**F**), defined in this example as

$$\mathbf{F} = \begin{pmatrix} |1 & 0 & 0 & 0| \\ |0 & 1 & 0 & 0| \\ |0 & 0 & 1 & 0| \end{pmatrix} \tag{4b}$$

Again we can verify that there are exactly three observed variables in the diagram, and the placement of the unit values in this matrix is defined by the order they were placed in the **p** vector.

practically useful as a tool because they can be used to represent the input and output of any of the SEM computer programs. This is not true when people create their own notation, such as including (a) unit vectors without two-headed arrows, (b) multiple unit vectors, or (c) arrows into other arrows. In such cases, we are left to consider what this all means. But the SEM path diagrams do not only substitute for the SEM algebraic interpretations, so typically the path diagram is used as a conceptual device. Examples where this happens will be pointed out.