In A. Madeau & J.J. McArdle, (Eds.). Contemporary Advances
in Psychometrics. Mahwah, NJ: Lawrence Erlbaum
Associates, Inc. (pp.225-273)

III

Structural Equation Models and Related Methods

Many of the seminal contributions of Roderick P. McDonald come from his work in structural equation modeling (SEM). His early contributions in this area (McDonald, 1969; McDonald & Swaminathan, 1972) were brought to the forefront in his covariance structure analysis (COSAN) model (McDonald, 1978) and the freely available computer program (developed with Colin Fraser). The COSAN model is known for its great flexibility, and reflects McDonald's deep understanding of the mathematical and statistical basis of SEM. Extensions of this logic are presented both in McDonald (1985) and in his Psychometric Society presidential address (McDonald, 1986). In more recent work, Rod has extended SEM in many ways, including theoretical work with Jack McArdle on RAM notation (McArdle & McDonald, 1984) and with Harvey Goldstein on multilevel SEM (McDonald & Goldstein, 1989), and in his continuing efforts to make SEM more widely available and understandable (e.g., McDonald, 1997, 2002).

This part of the book starts with a chapter by **Jack McArdle** describing the initial development of RAM rules for latent variable structural equation modeling. This chapter discusses the interplay between general SEM models (e.g., LISREL and COSAN), as well as the important personal contributions of Rod McDonald to progress in this area.

The next three chapters are devoted to important statistical issues in SEM. The broad problem of goodness of fit in SEM is an important topic taken up by **Herbert Marsh**, **Kit-Tai Hau**, and **David Grayson**. This is a description of contributions of many behavioral scientists to this vexing problem, and Rod McDonald's work is clear and evident. **Wolfgang Hartmann**, known as the creator of SAS PROC CALIS, discusses the practical and theoretical basis of resampling methods in SEM. This chapter leaves no doubt about the important future of this topic.

Jim Steiger details an important and often overlooked problem in basic SEM statistical tests—the comparison of correlations between and within independent samples. Steiger shows why this problem is initially difficult and how it can be easily solved using SEM.

The last two chapters of this part expand the current SEM in important directions. **Michael Browne** and **John Nesselroade** discuss the represention of psychological processes with dynamic factor models based on ARMA time-series models. **Harvey Goldstein** and **William Browne** discuss multilevel factor analysis models and show how they can be used with both continuous and discrete data. Both chapters are filled with interesting modeling ideas that will be widely used in future SEM research.

REFERENCES

McArdle, J. J., & McDonald, R. P. (1984). Some algebraic properties of the reticular action model for moment structures. *British Journal of Mathematical and Statistical Psychology, 37*. 234–251.

McDonald, R. P. (1969). A generalized common factor analysis based on residual covariance matrices of prescribed structure. *British Journal of Mathematical Statistical Psychology,* 22, 149–163.

McDonald, R. P. (1978). A simple comprehensive model for the analysis of covariance structures. *British Journal of Mathematical and Statistical Psychology, 31*, 59–72.

McDonald, R. P. (1985). *Factor analysis and related methods.* Hillsdale, NJ: Lawrence Erlbaum Associates, Inc.

McDonald, R. P. (1986). Describing the elephant: Structure and function in multivariate data. *Psychometrika, 51*, 513–534.

McDonald, R. P. (1997). Haldane's lungs: A case study in path analysis. *Multivariate Behavioral Research, 32*, 1–38.

McDonald, R. P. (2002). What can we learn from the path equations? Identifiability, constraints, equivalence. *Psychometrika, 67,* 225–249.

McDonald, R. P., & Goldstein, H. (1989). Balanced versus unbalanced designs for linear structural relations in two-level data. *British Journal of Mathematical and Statistical Psychology, 42*, 215–232.

McDonald, R. P., Ho, M.-H., & Ringo, A. F. (2002). Principles and practice in reporting structural equation analyses. *Psychological Methods, 7*, 64–82.

McDonald, R. P., & Swaminathan, H. (1972). *Structural analyses of dispersion matrices based on a very general model with a rapidly convergent procedure for the evaluation of parameters.* Toronto: University of Toronto, Ontario Institute for Studies in Education, Department of Measurement, Evaluation and Computer Applications.

9

The Development of the RAM Rules for Latent Variable Structural Equation Modeling

John J. McArdle
University of Virginia

This chapter describes the original basis of the Reticular Action Model (RAM) approach to structural equation modeling (SEM). The RAM rules were initially presented (1978–1981) to simplify and unify structural equation models based on path analysis graphics. The mathematical representation of RAM is presented as a second-order moment structure model that includes two parameter matrices, one of which is a patterned inverse. The graphic representation of RAM is presented through a series of definitions and axioms that provide a complete and concise isomorphism between graphics and algebra. Comparisons to other traditional models, such as those of multiple linear regression, path analysis, and factor analysis, show how these models may be easily and economically represented using RAM rules. Two other general modeling foundations, LISREL and COSAN, are presented as special cases and, somewhat paradoxically, as generalizations of RAM. These results are then used to develop some important technical features of the RAM rules, including efficient algorithmic estimation procedures and the further development of statistical indicators. Issues of conceptual representation are provided from a general systems perspective. Finally, some of the current limitations and benefits of the RAM rules are considered.

INTRODUCTION

Researchers interested in structural equation modeling often face the same questions that were apparent 25 years ago: (a) What analyses can I do with SEM? (b) How do I do these SEM analyses? (c) Why should I do these SEM analyses? Various answers to these questions have been refined over the last several decades, and expert treatments of these topics are available (e.g., Loehlin, 1987, 1998; McDonald, 1986, 1995). However, a substantial part of the debate and confusion about these key questions emanates from the fundamental frame used to describe SEM. To deal with these questions, I revive ideas surrounding what were termed "general models" for SEM. In the early 1970s the question of SEM generality was a major issue for both SEM developers and users. Two seminal developments along these lines included the creation of the linear structural relations (LISREL) program by Jöreskog (1973b) and the creation of the covariance structure analysis (COSAN) program by McDonald (1978). I used these classical ideas to develop a simple notation termed the reticular action meta-model (RAM; McArdle & McDonald, 1984) that renders some of these controversies and problems moot.

The SEM problems discussed here often arise when we deal with practical problems, for example, when we use SEM with longitudinal data with more than two occasions, with any models including incomplete data patterns, and with any models based on atypical representations of latent variables (e.g., Horn & McArdle, 1980; McArdle, 2001). A recent model including this kind of complexity is the latent variable SEM presented in Fig. 9.1; this is the precise path-graph representation of a bivariate difference score model for longitudinal twin data (McArdle & Hamagami, 2003). Of course, any sensible researcher may ask, "Why would anyone want to fit such a model?" Although this is a reasonable question, this is not the focal topic here. Instead we presume a model of this complexity is of substantive interest. This leads the researcher to ask, "Can this model be fitted with my favorite SEM program?" If so, "How can we actually fit this model?" One purpose of this chapter is to provide practical answers to questions of this nature. To do this I will show that the RAM rules are based on the *necessary and sufficient* features required for a general SEM notation.

To begin this task, it is useful to recognize that many researchers rely on the dominant SEM notation termed linear structural relations (LISREL; e.g., Jöreskog & Sörbom, 1979). This elegant notation was developed by merging concepts from a psychometric common factor "measurement model" combined with an econometric "structural model" (Jöreskog, 1973b; Wiley, 1973). In this form LISREL creates restrictive hypotheses about the manifest covariance structure, so this two-part distinction is often seen as a fundamental SEM requirement (e.g., Bollen, 1989; Hayduk, 1987). On the surface, this also seems to imply that LISREL cannot deal with some important analytic problems: (a) path models with no latent variables, (b) correlations among the specific factors across blocks of variables, (c) estimation of standard deviations rather than variances, (d) direct arrows from manifest to

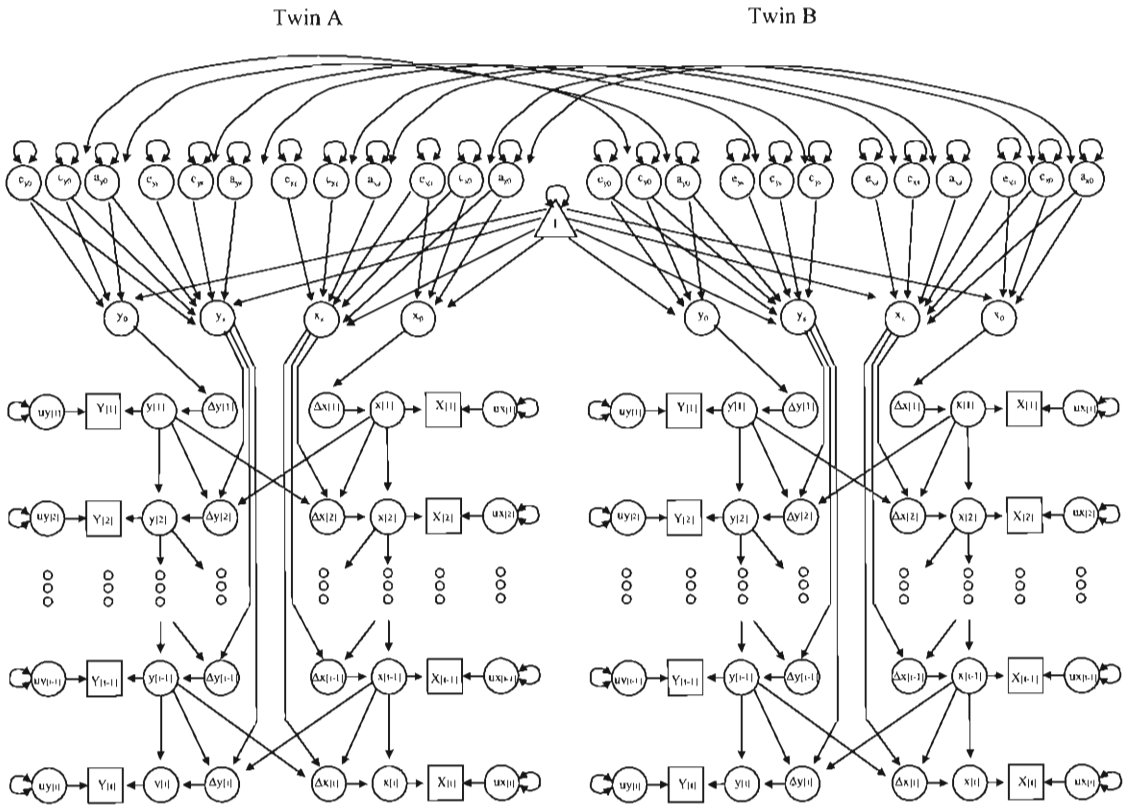

FIG. 9.1. An example of a contemporary path diagram. *Note*: From *Structural equation models for evaluating dynamic concepts within longitudinal twin analyses* by J. J. McArdle & F. Hamagami (2003). *Behavior Genetics, 33*(3), 137–159. Copyright 2003 by Springer Science+Business Media, Inc. Reprinted with permission.

latent variables, (e) higher order factors within each set, (f) more than two sets of variables (e.g., X, Y, and Z), (g) means and intercept terms in the equations, and (h) allowing correlations between predictors and disturbance terms. However, in an illuminating series of reports, Jöreskog and Sörbom (1979) demonstrated how the LISREL program could be reparametrized, or "tricked," into handling these problems. These SEM possibilities still create confusion among the expert and novice alike, but the flexibility of the computational algorithm (combined with the use of the LISREL **B** matrix) leads to an interesting result: *The LISREL model is a general organizing principle but actually poses no restrictive hypotheses about any covariance structure.*

At about the same time, McDonald (1978) developed an algebraic model and implemented a computer program, termed COSAN, which used a higher order structure that seemed to be flexible enough to carry out all SEM analyses. McDonald (1978, p. 61) basically showed that symmetric matrices at each level of a latent variable higher order system were not needed. Any symmetric matrix could essentially be "carried up" from one level to the next by using a series of latent variables with no variance, that is, fixed linear combinations. This mathematical simplicity led to a series of elegant statistical features described by McDonald (1978, 1979, 1980; McDonald & Swaminathan, 1972). This use of flexible programming in place of substantive concepts made COSAN much different from the basic concepts of the prior models (Bentler, 1976).

As a postgraduate user of both programs, I found myself spending a great deal of time studying the ways that these programs were dealt with both standard models and special cases. I learned the differences in the input scripts required for these programs. This allowed me to represent the output of both programs as path diagrams. It was at this output stage that I noticed the communalities among all such programs. I found I could obtain exactly the same path analysis results from both programs. The special cases I was studying were precursors to models such as those of Fig. 9.1, and that did not seem to fit easily into either program (e.g., Horn & McArdle, 1980). The success of these analyses led me to the unusual conclusion that *there were no real differences between LISREL and COSAN.*

My initial publications were suggestions to use a slightly new set of rules for complete path analysis diagrams (McArdle, 1978, 1979b, 1979c). However, this approach also highlighted an algebraic paradox: These graphics showed not only that LISREL and COSAN were identical in applications, but that they also had the same algebraic limits, and each could do the same analyses with far less matrix algebra. In essence, *LISREL and COSAN were special cases of each other and themselves.* This suggestion contested most previous work on this topic, and these paradoxical ideas were dealt with in different ways by different researchers. In general, this was not a critical issue to most SEM researchers because both LISREL and COSAN worked exceptionally well. However, it was also clear that that any advances in this area were thought to be found in extensions of LISREL and COSAN. For example, Bentler (1976) proposed an extension of COSAN as

a higher order common factor model, and his student Weeks (1978) completed a doctoral dissertation on an entirely new program based on this logic. In a public presentation (McArdle, 1978), I demonstrated how higher order factor models could already be programmed in both LISREL and COSAN, and this meant that new computer programs were not necessary for this or other problems. In a series of subsequent presentations, Bentler and Weeks (1979, 1980) altered the previous general representation of Bentler (1976) in a radical fashion to account for this new finding.

At this time one of the key developers in this field, Roderick P. McDonald (who I had not yet met) tried to understand my paradoxical conclusions. Rod's willingness to work with me was most helpful in the early days of the basic RAM rules. Under Rod's guidance, I improved the original 1978 manuscript and circulated it in 1981 under the title, "The Reticular Action Meta-model: A Simplified Representation for Structural Equation Modeling." This manuscript was difficult to publish, partly due to the confusing paradoxes and the controversy over its development. After several years of rejection, the key algebraic aspects of this RAM presentation were published in McArdle and McDonald (1984). Some key graphic principles of RAM were published in the RAMpath program by McArdle and Boker (1990) and Boker, McArdle, and Neale (2002). The original and complete development of RAM is published here for the first time as a small tribute to Roderick P. McDonald.

To preserve some key historical aspects of this treatment, the next six sections appear in the exact form as the initial presentation of March 1981. An inital historical account of research on general SEM concepts (ca. 1981) is followed by a description of the mathematical basis of RAM notation (with Greek notation where appropriate), a description of the graphic basis of RAM notation as derived from the mathematics (with highlights where appropriate), and a comparison of RAM notation to several alternatives, which leads to an exposition of the paradoxical interrelationships among general systems, especially LISREL and COSAN, in both graphic and algebraic terms. Then some computational features are added to suggest how RAM notation could be used as the generic basis of future computer programs. This is followed by a discussion of the conceptual utility of RAM, including the freedom to use SEMs that do not seem to fit into any more specific schema (e.g., Fig. 9.1). In the concluding section, I offer an overview of the first 25 years of the RAM rules.

HISTORICAL INTRODUCTION (CIRCA 1981)

There has recently been some interest in the relationships among models that have been proposed for the structural analysis of covariance matrices and mean vectors (i.e., moment structures). Not surprisingly, in the chronological sequence of

development in this area one seems to find a progression from special to more general models that subsume their predecessors. Broadly, the early sequence is from Spearman's (1904) classical one-factor model, through Thurstone's (1947) multiple-factor model, to the beginnings of the treatment of the restricted oblique common factor model by Anderson and Rubin (1956). These seminal developments led to Bock and Bargmann's (1966) introduction of a group of models for the analysis of covariance structures, which was later expanded into the general analysis of covariance structures (ACOVS) model of Jöreskog (1970). The ACOVS model is essentially a second-order common factor model in which the matrices may be *patterned*; that is, the elements of any matrix could be prescribed constants, parameters that are free to be individually estimated, or parameters that are constrained to be equal to one or more other parameters. These provisions supplied a remarkably flexible model for the analysis of covariance structures (Jöreskog, 1973a).

There have been several more recent attempts (*sic*) to develop these ideas further. For example, Bentler (1976) recommended the development of a straightforward generalization of ACOVS to what might be called an mth-order common factor model. However, neither Jöreskog's ACOVS model nor Bentler's mth-order counterpart of it made provision for specifying the parameters of the inverse of a matrix in the model. Subsequently, Keesling and Wiley (see Wiley, 1973) showed a patterned inverse was necessary in a general-purpose model for multivariate analysis and Jöreskog (1973b) used this in his model for linear structural relations (LISREL). We may loosely describe LISREL as a third-order restricted factor model in which the inverse of one of the three factor loading matrices is patterned. Because a patterned matrix does not in general have an inverse of the same pattern, specific provision for patterned inverses is thus seen to be necessary in the analysis of moment structures. Jöreskog and Sörbom (1980) programmed a version of this model (LISREL-IV) that allowed patterned residual covariances, as in McDonald (1969), and multiple independent groups, as in Jöreskog (1971).

McDonald (1978) described an mth-order model in which the inverse of any matrix of the model can be patterned as desired. As a consequence, McDonald's model, COSAN, yields the ACOVS model of any order when no inverse matrix is patterned, and also models of the LISREL type with a mixed sequence of patterned matrices and patterned inverses. McDonald obtained some simplification of the mth-order model by an algebraic device that allowed it to be written as though it contained no residual covariance matrices. This meant that an mth-order factor model would contain only $m + 1$ matrices instead of $2m + 1$, thus providing a remarkable economy of representation that simplified the programming of the model and certain of its properties.

In a general sense, then, the chronological sequence that has just been outlined shows a broad tendency toward increasing complexity of the matrix representation of the models accompanying an apparent increase in generality. In direct contrast, however, McArdle (1978, 1979b, 1979c, 1980; Horn & McArdle, 1980) presented

a path-graphic formulation termed RAM, which showed that all of the foregoing models for the structural analysis of multivariate data could be represented extremely economically by using only a second-order model with an inverse feature. This model also demonstrated a paradoxical set of interrelationships among the models and rendered meaningless any question of determining which model is a "special case" of which other from their matrix specification alone.

A specialized version of McArdle's general result was quoted without proof by Bentler and Weeks (1979, Equations 5 and 6). The account given by Bentler and Weeks of this representation is far from complete and at several points might be misleading. For example, Bentler and Weeks dismiss the simple RAM representation as being "not as convenient for comparative purposes" (p. 172) as, for example, the Bentler (1976) *m*th-order representation. Thus, the object of this chapter is to fully develop the RAM representation of multivariate moment structures in a formal fashion at several levels of representation; including mathematical, graphic, comparative, technical, and conceptual. This discussion will serve to both clarify the relations between RAM and the main models in this field and show how the simple principles of RAM provide a *necessary and sufficient foundation for traditional* structural equations models.

ALGEBRAIC REPRESENTATION

Let $\mathbf{v}$ be a $t \times 1$ vector of random *variables*, and let

$$\mathbf{v} = \mathbf{A}\,\mathbf{v} + \mathbf{u}, \tag{1}$$

where $\mathbf{A}$ is a square matrix of structural weights of $\mathbf{v}$ on itself so that any element α_{ij} represents a *directed* or *asymmetric* relationship from column variable $\mathbf{v}_j$ to row variable $\mathbf{v}_i$. In the typical case where a variable is not directed onto itself, Equation 1 is similar to the basic regression equation of image theory as treated by Guttman (1957); $\mathbf{A}$ represents the regression of each of the t variables on the other $t - 1$ variables; $\mathbf{u}_i$ represents the residual or anti-image of $\mathbf{v}_i$; and the corresponding diagonal α_{ii} is zero. If all regression coefficients on other variables are zero (e.g., $\alpha_{ij} = 0$, $j = 1, \ldots, t$), then the variable $\mathbf{v}_i$ is considered the same as its own residual $\mathbf{u}_i$.

We also define the $t \times t$ matrix

$$\mathbf{\Omega} = E\{\mathbf{u}\,\mathbf{u}'\}, \tag{2}$$

where E denotes the expectation operator, so that any element ω_{ij} represents an *undirected* or *symmetric* structural relation among residual variables $\mathbf{u}_i$ and $\mathbf{u}_j$. If there is no relationship among the $\mathbf{u}_i$ and $\mathbf{u}_j$, it follows that $\omega_{ij} = 0$ ($i = 1, \ldots, t$; $j = 1, \ldots, t$).

The structural parameters of $\mathbf{A}$ and $\mathbf{\Omega}$ are presumed to be associated with a set of numerical values. In terms of model specification, we will permit any of these parameters to take on one of three different settings: (a) *fixed* at some prescribed constant (e.g., zero or unity), (b) *free* to be estimated in some prescribed fashion, and/or (c) *constrained* to be a linear function of one or a set of other parameters (e.g., equal to another). More details on the estimation of these values are discussed in a later section.

To define the overall moments among variables in terms of structural parameters, we define the $t \times t$ symmetric matrix

$$\mathbf{\Sigma} = E\{\mathbf{v}\,\mathbf{v}'\}, \tag{3}$$

where each element σ_{ij} represents the overall *association* between variables $\mathbf{v}_i$ and $\mathbf{v}_j$. Using these basic definitions, it follows that Equation 1 can be rewritten as

$$\mathbf{u} = \mathbf{v} - \mathbf{A}\mathbf{v} = (\mathbf{I}_t - \mathbf{A})\mathbf{v}, \tag{4}$$

where $\mathbf{I}_t$ is a $t \times t$ identity matrix, whence it follows that

$$\mathbf{\Omega} = (\mathbf{I}_t - \mathbf{A})\mathbf{\Sigma}(\mathbf{I}_t - \mathbf{A})'. \tag{5}$$

Perhaps more important, we may now also write all variables in terms of the residual structural relations as

$$\mathbf{v} = (\mathbf{I}_t - \mathbf{A})^{-1}\mathbf{u}, \tag{6}$$

which yields the *moment structure model* for all variables as

$$\mathbf{\Sigma} = (\mathbf{I}_t - \mathbf{A})^{-1}\mathbf{\Omega}(\mathbf{I}_t - \mathbf{A})^{-1'}. \tag{7}$$

For purposes of empirical utility, we now let $\mathbf{v}$ be partitioned into two subvectors $\mathbf{m}$, of j components, and ℓ, of k components. That is, $t = j + k$ and

$$\mathbf{v}' = [\mathbf{m}, \ell]'. \tag{8}$$

The j components of $\mathbf{m}$ may be considered observed or *manifest* variables, and the k components of ℓ may be considered unobserved or *latent* variables. To distinguish these variables we define the $j \times t$ matrix

$$\mathbf{F} = [\mathbf{I}_j : {}_j\mathbf{O}_k], \tag{9}$$

where ${}_j\mathbf{O}_k$ is a $j \times k$ null matrix, so that

$$\mathbf{m} = \mathbf{F}\mathbf{v}. \tag{10}$$

Thus, $\mathbf{F}$ is a fixed known matrix of prescribed unity and zero constants that acts to select or *filter* the manifest variables out of the full set of manifest and latent variables. If, for any reason, the components of $\mathbf{v}$ are permuted to some mixed order, the columns of $\mathbf{F}$ can be correspondingly permuted.

We now write the $j \times j$ symmetric matrix

$$\boldsymbol{\Sigma}_{mm} = E\{\mathbf{m}\mathbf{m}'\} \tag{11}$$

as a general raw product–moment matrix without correction for means. Of course, with a choice of scale such that all manifest variables have zero means, $\boldsymbol{\Sigma}_{mm}$ will be a covariance matrix. From Equations 5 and 12 we may also write

$$\mathbf{m} = \mathbf{F}(\mathbf{I}_t - \mathbf{A})^{-1}\mathbf{u}, \tag{12}$$

whereby we may finally write the moment structure among manifest variables as

$$\boldsymbol{\Sigma}_{mm} = \mathbf{F}\boldsymbol{\Sigma}\mathbf{F}' = \mathbf{F}(\mathbf{I}_t - \mathbf{A})^{-1}\boldsymbol{\Omega}(\mathbf{I}_t - \mathbf{A})'^{-1}\mathbf{F}'. \tag{13}$$

The resulting equation completely specifies the expectations for the manifest variable moments $\boldsymbol{\Sigma}_{mm}$ in terms of the fundamental RAM structural parameters $\mathbf{A}$ and $\boldsymbol{\Omega}$.

In most applications the model matrices for RAM are usually somewhat large and sparse when compared with the traditional algebraic forms. This will generally be true for parameter matrices $\mathbf{A}$ and $\boldsymbol{\Omega}$ as well as resultant matrices $(\mathbf{I} - \mathbf{A})$ and $\boldsymbol{\Sigma}$. However, to represent all features of any model with sparse matrices we need only (a) provide the relevant manifest j and latent k matrix dimensions, (b) reorganize any long column vectors into transposed row vectors, and (c) use block-partitioned row and column subscripts to refer only to the nonzero subpartitions of $\mathbf{A}$, $\boldsymbol{\Omega}$, and resultant $\boldsymbol{\Sigma}$.

The matrix representations of Equations 6 and 7 make the assumption that matrix $\mathbf{I} - \mathbf{A}$ is nonsingular, and it may be useful to more explicitly define

$$\begin{aligned}\mathbf{E} = (\mathbf{I} - \mathbf{A})^{-1} &= \sum_{i=0}^{\infty} \mathbf{A}^i = \mathbf{I} + \mathbf{A}^1 + \mathbf{A}^2 + \cdots \\ &= \prod_{i=0}^{\infty} \mathbf{A}^i = [\mathbf{Cof}(\mathbf{I} - \mathbf{A})]|\mathbf{I} - \mathbf{A}|^{-1},\end{aligned} \tag{14}$$

where $\mathbf{E}$ represents a matrix of total effects and $\mathbf{Cof}(\cdot)$ represents the matrix of cofactors. These inverse representations, which may all be verified through multiplication by $(\mathbf{I} - \mathbf{A})$, prove particularly useful in the elemental structural decomposition of the association matrix $\boldsymbol{\Sigma}$. Following Bock (1975), we may write the determinant of any square matrix as the algebraic sum of all possible products of elements in the matrix; each product is formed in such a way so that one and only

one element from each row and column appears in the product. In the specific case of Equation 14, we may expand the determinant as the elemental products

$$|\mathrm{E}| = |\mathbf{I} - \mathbf{A}| = \sum_{i=1}^{t^2!} (-1)^s \varepsilon_{1x} \varepsilon_{2y} \cdots \varepsilon_{(\mathrm{t}\,\mathrm{z})}^2, \tag{15}$$

where $\varepsilon_{ij} = (-\alpha_{ij})$ for $i \neq j$ or $(1 - \alpha_{ij})$ for $i = j$; $s = +1$ or -1 depending on whether the natural permutation order of the row subscripts is odd or even, respectively; and $x, y, \ldots, z$ represent the permutations of natural order from 1 to t. The element in the *kth* row and *jth* column of the cofactor matrix in Equation 8 may now be written as the determinant of a minor matrix whose jth row and kth column have been removed. These algebraic relationships show that any specific element σ_{ij} of the overall association matrix $\boldsymbol{\Sigma}$, when calculated as the bilinear form of Equation 7, may be decomposed into (a) a row sum of signed products based on ε_{ig}, (b) the multiplication of an element $\mathbf{u}_{gh}$, and (c) the postmultiplication by a transposed row sum of signed products based on ε_{hj}. (Of course, as Equation 14 shows, if $|\mathbf{I} - \mathbf{A}| \neq 1$, these cofactors must all be scaled accordingly.) For purposes of *exposition as path-graphic tracing rules*, we later write the resultant algebra of $\boldsymbol{\Sigma}$ in this full determinant–product decomposition. Specific examples using this more compact notation are presented when useful.

GRAPHIC REPRESENTATION

A critically important aspect in the development of RAM theory is the precise one-to-one relationship between the algebraic form just described and a formal graphic system. Some of our earlier work (Horn & McArdle, 1980; McArdle, 1978, 1980) emphasized the development of the algebra as a direct result of the graphics. Let us first present the properties of any general structural system by offering a series of definitions that describe its primitive features, the axioms that relate the algebra directly to the graphics, and a set of figures that illustrate our major points.

Definition 1. A finite collection of *t points* or vertices or nodes in any space is contained in set **v**.

Axiom 1. Points may be used to represent variables **v**. Figure 9.2a illustrates this use of variables as points. These are circled and labeled $\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_t$.

Definition 2. A finite collection of *z lines* or edges that connect each pair of points is contained in set $\boldsymbol{\Sigma}$.

Axiom 2. Lines between points may be used to represent the associations between variables in matrix $\boldsymbol{\Sigma}$. Figure 9.2b presents an example of these as associations as lines, and shows the line between points $\mathbf{v}_i$ and $\mathbf{v}_j$ labeled as σ_{ij}.

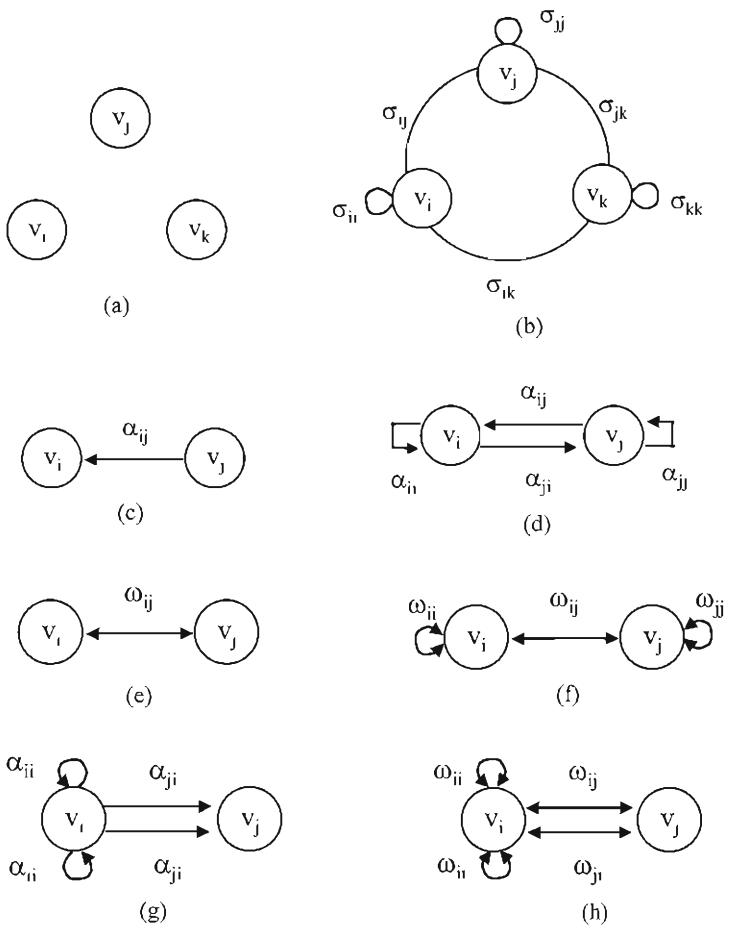

FIG. 9.2. Basic graphic elements of the reticular action meta-model. (a) Variables as points in space. (b) associations as lines between points, (c) a directed arrow, (d) a full set of directed arrows, (e) an undirected turn, (f) a full set of undirected turns, (g) inconsistent arrows, (h) inconsistent turns.

Definition 3. A collection of x directed arrows between pairs of points is contained in set **A**. In any pair of points, one point will be the *input*, or starting point of the directed arrow, and the other will be the *output*, or ending point.

Axiom 3. Directed arrows from one point to another may be used to represent the coefficients of matrix **A**. For any arrow, the input variable is indexed in the column of **A** and the output variable is indexed in the row of **A**. Figure 9.2c presents a simple case of this kind, which shows that the most crucial graphic aspects of

an arrow are its direction and value. Figure 9.2d is slightly more complex because it illustrates that a directed arrow may be simultaneously drawn between any two edges of any point, including the possibility of a reflexive "self-loop" arrow from a point back to itself (Heise, 1975). *This self-loop is a novel feature of path graphics that emerges from the algebraic definitions.*

Definition 4. A collection of y undirected *turns* or *slings* between pairs of points is contained in set $\boldsymbol{\Omega}$.

Axiom 4. Undirected slings between points may be used to represent the elements in matrix $\boldsymbol{\Omega}$. In any specific sling, both variables are indexed by the respective row and column entries of $\boldsymbol{\Omega}$. Figure 9.2e illustrates that the undirected sling between the pair of points $\mathbf{v}_i$ and $\mathbf{v}_j$ is drawn as a line with two arrowheads and labeled as ω_{ij}. Figure 9.2f illustrates that the complete set of slings can include the possibility of a "self-sling" with itself (i.e., ω_{ii} and ω_{jj}) as well as with other points (i.e., ω_{ij}). *This self-sling is a novel feature of path graphics that emerges from the algebraic definitions.*

Definition 5. Parallel elements are pairs of lines, pairs of arrows, or pairs of turns that span the same pair of points and, in the case of arrows, point in the same direction. Only *nonparallel* elements may be contained within set $\boldsymbol{\Sigma}$, $\mathbf{A}$, or $\boldsymbol{\Omega}$.

Axiom 5. Parallel arrows, turns, and lines are not permitted because only one value can be placed in any single row and column location of the matrix $\boldsymbol{\Sigma}$, $\mathbf{A}$, or $\boldsymbol{\Omega}$. Figure 9.2g illustrates the inconsistent condition of two parallel sets of arrows, those for α_{ji} and those for α_{ii}. Figure 9.2h illustrates the inconsistent condition of two parallel turns, those for ω_{ij} and those for ω_{ii}. Alternatively, we note that the condition where $\alpha_{ij} \neq 0$ and $\omega_{ij} \neq 0$ does not represent a parallelism and may be consistent.

Definition 6. Every element in set $\mathbf{A}$ and $\boldsymbol{\Omega}$ may take on a symbolic or numeric *value*. Under the condition that the specific value is symbolically null or numerically zero, that specific element will not be drawn. Under the condition that the specific value is symbolically or numerically unit valued, that specific element will not be labeled.

Axiom 6. The value of any line, arrow, or turn may be entered, either symbolically or numerically, in the corresponding row and column location of $\boldsymbol{\Sigma}$, $\mathbf{A}$, or $\boldsymbol{\Omega}$. All elements that are not drawn are assumed to be null.

Definition 7. A *graph* is a collection of t points $\mathbf{v}$, x directed arrows $\mathbf{A}$, and y undirected slings $\boldsymbol{\Omega}$, whose values may be used to determine the corresponding values for the z lines in $\boldsymbol{\Sigma}$.

Axiom 7. A graph is a precise representation of a structural *model*. The value of the arrows in $\mathbf{A}$ and the turns in $\boldsymbol{\Omega}$ may be used to provide a "structure" for

the values in $\boldsymbol{\Sigma}$ according to Equation 7. For this reason, we may now say that the arrows in $\mathbf{A}$ and the turns in $\boldsymbol{\Omega}$ provide a "structure" for the lines in $\boldsymbol{\Sigma}$. It also follows from these definitions that any graph is composed of nonzero values for $0 \leq x \leq t^2$ arrows in $\mathbf{A}$ and $0 \leq y \leq t(t+1)/2$ turns in $\boldsymbol{\Omega}$, which are used to determine the values of the $0 \leq z \leq t(t+1)/2$ nonzero lines in $\boldsymbol{\Sigma}$.

Definition 8. A *pathway*, termed $\pi(ij)$, in any graph is any collection of b backward-directed arrows, one turn, and f forward directed arrows, which start at one point $\mathbf{v}_i$ and end at one point $\mathbf{v}_j$. The value of a pathway $\pi(ij)$ is the *product* of the value of its associated $\mathbf{A}$ and $\boldsymbol{\Omega}$ elements.

Axiom 8. A pathway $\pi(ij)$ may be used to represent the product of (a) a single b-termed element of $\mathbf{E}$, (b) a single element of $\boldsymbol{\Omega}$, and (c) a single f-termed element of $\mathbf{E}'$. The number of product terms required to make up each specific element of $\mathbf{E}$ is based on the nonzero configuration of elements in $\mathbf{A}$.

Figure 9.3a presents the general form of a pathway $\pi(i \ldots j)$ from point $\mathbf{v}_i$ to point $\mathbf{v}_j$. This pathway is initiated from point $\mathbf{v}_i$ and flows backward along a set of b directed arrows. These arrows must all point backward, connecting $b+1$ distinct points, and terminate at some point $\mathbf{v}_g$. At point $\mathbf{v}_g$, a single undirected turn toward a forward flow is made through a two-headed ω_{gh} between $\mathbf{v}_g$ and $\mathbf{v}_h$. This single turn from backward to forward arrowhead direction is the only consistent way to alter the pathway direction. Flow then proceeds from $\mathbf{v}_h$ along a set of f forward-directed arrows, passing through $f+1$ total points, until final destination point $\mathbf{v}_j$ is reached. It may be useful to label this general pathway as $\pi(ik \ldots gh' \ldots \ell' j')$, where the transpose is used to designate forward flow. *The definition of a pathway "tracing rule" emerges as a direct interpretation of the algebraic decomposition* (Equations 7 and 14).

This general expression now permits the statement of any specific pathway through a reorientation of pathway points, and this highlights unusual pathways. For example, by letting $i = j$ in the foregoing general expression, we find that the initial starting point is also the final end point and the total pathway is considered a "circuit." A specific example of this type is presented as pathway $\pi(123'1')$ in Fig. 9.3b. As another example, we now let $i = g$ and $b = 0$, so the pathway starts backward with a turn between variables 1 and 2. This case is depicted by $\pi(12'3')$ in Fig. 9.3c. Of course, the same logic would apply to a model where $h = j$ and $f = 0$, where pathway $\pi(12'3')$ would end on the forward flow of a turn. Figure 9.3d illustrates the case where $g = h$ and the self-turn is used as the turning point in $\pi(122'3')$. Figure 9.3e assumes $i = j$, $g = h$, and $b = f = 0$, so the entire pathway reduces to a self-turn denoted by $\pi(11')$. As an example of the pathway value calculation, we simply note that the value of $\pi(123'1')$, as presented in Fig. 9.3b, would be symbolically given as the three-term product $\alpha_{12}\omega_{23}\alpha'_{31}$. Alternatively, the value of $\pi(12'3')$ of Fig. 9.3c is given as the two-term product $\omega_{12}\alpha'_{23}$.

Definition 9. Two pathways are parallel if they contain an identical set of $\mathbf{A}$ arrows and $\boldsymbol{\Omega}$ slings in an identical order. This implies that all points are identical also.

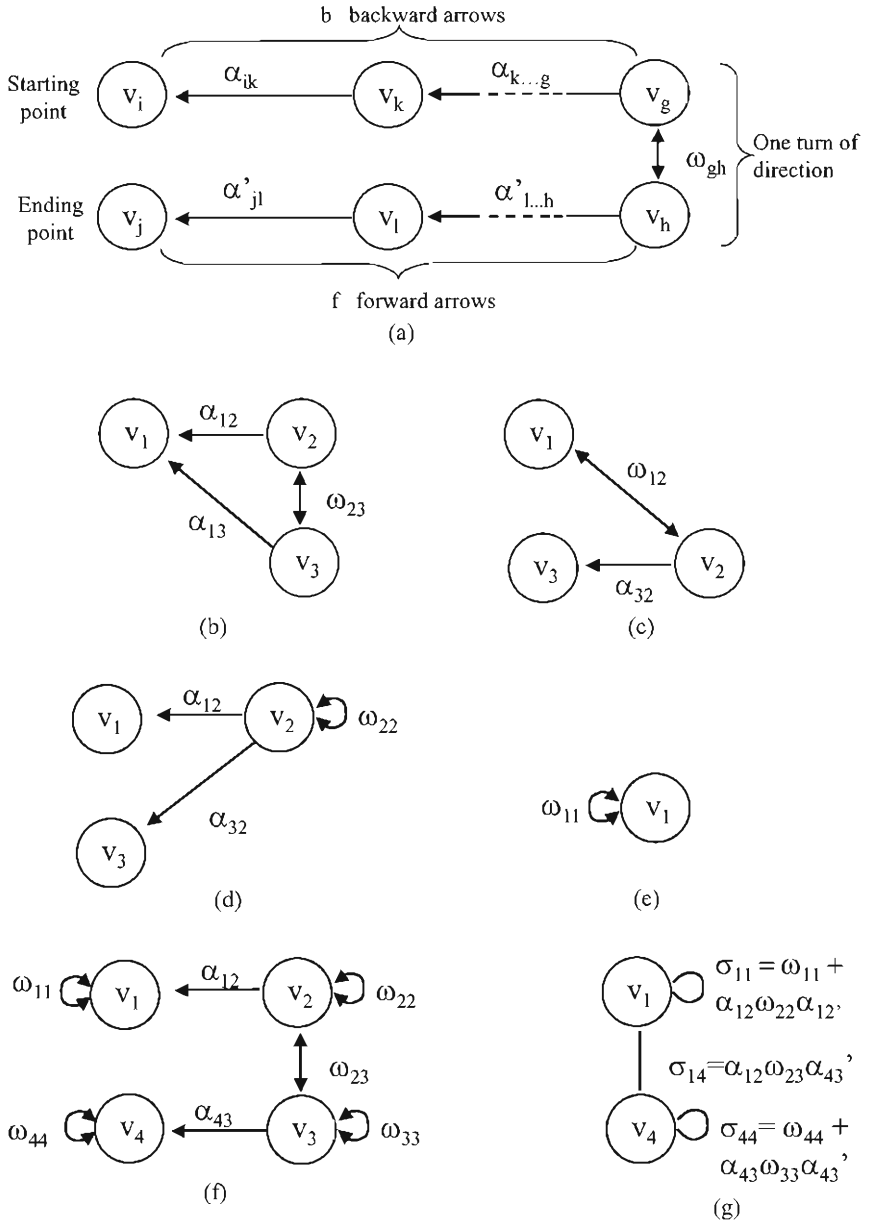

FIG. 9.3. Decomposition of association via pathway representation. (a) General expression of a single pathway $\pi(i..j)$, (b) a circuit as a full pathway, (c) starting (or ending) on a turn.

The total value of any association between any starting point and any ending point is the total *sum* of the values of the *nonparallel* pathways between these two points.

Axiom 9. The line of association between row variables and column variables may be used to represent the overall sum that is created as the bilinear form of Equation 7. In the diagonal and the lower symmetric quadrant of **A**, the row variable is the starting point and the column variable is the ending point.

Figure 9.3f illustrates a simple model with multiple pathways with simple labels. In this figure there is only one consistent pathway between variables 1 and 4; namely the pathway product $\alpha_{12}\omega_{23}\alpha'_{43}$. However, there are two consistent pathways each for the circuits from variables 1 and 4, the two pathways ω_{11} and $\alpha_{12}\omega_{22}\alpha'_{12}$ for variable 1 and the pathways labeled ω_{44} and $\alpha_{43}\omega_{33}\alpha'_{43}$ for variable 4. The lines of total association σ_{ij} are formed as the sum of these pathway products and drawn as Fig. 9.3g.

Definition 10. Nonoverlapping subclasses of points may be distinguished by their representation as geometric figures—*squares* for the j points labeled $\mathbf{m}_j$, *circles* for the k points labeled ℓ_k, etc.

Axiom 10. The graphic distinction between squares and circles may be used to represent the presence of a one or zero in the matrix **F** (see Equation 11).

Figure 9.4 presents an overall summary of the key features of the RAM graphics. Structural relations among these points are denoted by lines connecting these

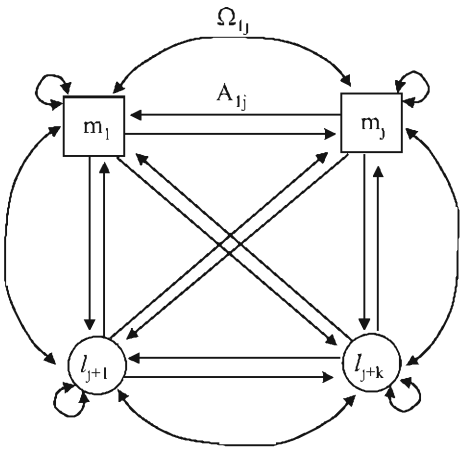

FIG. 9.4. An overview of reticular action meta-model graphics. Squares represent manifest variables and circles represent latent variables.

points, with directed arrows distinguished from undirected turns. The careful reader will note that this figure does not contain all possible structural relationships or the respective values. Whereas there exist many other graphic possibilities, only those explicitly described here are required for our further development.

As a brief example of the useful features of this isomorphism, consider that the model depicted in Fig. 9.3f may be simply translated into the RAM matrix algebra of

$$\boldsymbol{v} = \begin{bmatrix} v_1 \\ v_2 \\ v_3 \\ v_4 \end{bmatrix}, \quad \mathbf{A} = \begin{bmatrix} 0 & \alpha_{12} & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & \alpha_{43} & 0 \end{bmatrix} \quad \text{and}$$

$$\boldsymbol{u} = \begin{bmatrix} u_1 \\ v_2 \\ v_3 \\ u_4 \end{bmatrix}, \quad \boldsymbol{\Omega} = \begin{bmatrix} \omega_{11} & & & \text{sym.} \\ 0 & \omega_{22} & & \\ 0 & \omega_{23} & \omega_{33} & \\ 0 & 0 & 0 & \omega_{44} \end{bmatrix}, \tag{16}$$

where the elements of $\mathbf{u}$ are either labeled as residuals $\mathbf{u}_i$ or as the original variables $\mathbf{v}_i$ depending on the elements of $\mathbf{A}$ (i.e., if $\mathbf{v}_i$ is an outcome of any variable $\mathbf{v}_j$). We may now also write

$$\mathbf{E} = (\mathbf{I} - \mathbf{A})^{-1} = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix} + \begin{bmatrix} 0 & \alpha_{12} & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & \alpha_{43} & 0 \end{bmatrix} = \begin{bmatrix} 1 & \alpha_{12} & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & \alpha_{43} & 1 \end{bmatrix}, \tag{17}$$

which may be verified by $(\mathbf{I} - \mathbf{A})$ multiplication, so that

$$\begin{aligned} \boldsymbol{\Sigma} &= (\mathbf{I}_t - \mathbf{A})^{-1}\boldsymbol{\Omega}(\mathbf{I}_t - \mathbf{A})^{-1'} \\ &= \begin{bmatrix} \mathbf{1} & \alpha_{12} & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & \alpha_{43} & 1 \end{bmatrix} \begin{bmatrix} \omega_{11} & 0 & 0 & 0 \\ 0 & \omega_{22} & \omega_{32} & 0 \\ 0 & \omega_{23} & \omega_{33} & 0 \\ 0 & 0 & 0 & \omega_{44} \end{bmatrix} \begin{bmatrix} \mathbf{1} & 0 & 0 & 0 \\ \alpha_{12'} & 1 & 0 & 0 \\ 0 & 0 & 1 & \alpha_{43'} \\ 0 & 0 & 0 & 1 \end{bmatrix} \\ &= \begin{bmatrix} \omega_{11} + \alpha_{12}\omega_{22}\alpha_{12'} & & & \text{sym.} \\ \omega_{22}\alpha_{12'} & \omega_{22} & & \\ \omega_{23}\alpha_{12'} & \omega_{23} & \omega_{33} & \\ \alpha_{43}\omega_{32}\alpha_{12'} & \alpha_{43}\omega_{32} & \alpha_{43}\omega_{33} & \omega_{44} + \alpha_{43}\omega_{33}\alpha_{43'} \end{bmatrix} \end{aligned} \tag{18}$$

as illustrated in Fig. 9.3g. Although no distinction between circles and squares is made in Fig. 9.3f, any such distinction would merely require a subpartitioning of this $\boldsymbol{\Sigma}$ to obtain $\boldsymbol{\Sigma}_{mm}$. For example, under the assumption that $\mathbf{v}_1$ and $\mathbf{v}_4$ were

drawn in Fig. 9.3f as squares, we could write nonzero $\mathbf{f}_{11} = 1$ and $\mathbf{f}_{24} = 1$, and this would essentially filter out the associations given as Fig. 9.3g. This can be written in algebraic form as

$$\begin{aligned}\boldsymbol{\Sigma}_{\mathbf{mm}} &= \mathbf{F}\boldsymbol{\Sigma}\mathbf{F}' \\ &= \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} \omega_{11} + \alpha_{12}\omega_{22}\alpha_{12'} & & & \text{sym.} \\ \omega_{22}\alpha_{12'} & \omega_{22} & & \\ \omega_{23}\alpha_{12'} & \omega_{23} & \omega_{33} & \\ \alpha_{43}\omega_{32}\alpha_{12'} & \omega_{43}\omega_{32} & \alpha_{43}\omega_{33} & \alpha_{43}\omega_{32}\alpha_{12'} \end{bmatrix} \begin{bmatrix} 1 & 0 \\ 0 & 0 \\ 0 & 0 \\ 0 & 1 \end{bmatrix} \\ &= \begin{bmatrix} \omega_{11} + \alpha_{12}\omega_{22}\alpha_{12'} & \text{sym.} \\ \alpha_{43}\omega_{32}\alpha_{12'} & \alpha_{43}\omega_{32}\alpha_{12'} \end{bmatrix}\end{aligned} \tag{19}$$

Because we always employ the same matrix operations, we can create various simple devices to convey this information. For example, any graph may be even more simply translated into a compact RAM notation by including only the salient nonnull *list operator* $\mathbf{L}\{*\}$ where elements of Equations 16–19 are listed as

$$\begin{aligned} &\mathbf{L}\{\mathbf{v}\} = \{\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3, \mathbf{v}_4\}, \qquad \mathbf{L}\{\mathbf{A}\} = \{\alpha_{12}, \alpha_{43}\}, \\ &\mathbf{L}\{\mathbf{u}\} = \{\mathbf{u}_1, \mathbf{u}_2, \mathbf{u}_3, \mathbf{u}_4\}, \qquad \mathbf{L}\{\Omega\} = \{\omega_{11}, \omega_{22}, \omega_{33}, \omega_{44}, \omega_{34}\}, \\ &\mathbf{L}\{\mathbf{E}\} = \{1_{11}, 1_{22}, 1_{33}, 1_{44}, \alpha_{12}, \alpha_{43}\}, \qquad \mathbf{L}\{\mathbf{F}\} = \{1_{11}, 1_{44}\}. \end{aligned} \tag{20}$$

The resultant RAM algebra can always be summarized in list form as in Equations 20, and the elements of $\boldsymbol{\Sigma}$ in Equation 18 or elements of $\boldsymbol{\Sigma}_{mm}$ in Equation 19 could also be listed by row and column entries. These lists $\mathbf{L}$ may be most useful for computer input and output (as described in a later section).

This simple example illustrates the presence of at least one critical distinction between the algebra and the graphs. The inclusion of the self-loop and self-sling (see Figs. 9.2 and 9.3) illustrates that we have not provided a separate "point-in-space" representation for structural residual variables $\mathbf{u}$. As will be shown later, this representation turns out to be redundant in both graphic and algebraic terms, so, although possible, *it is not necessary to draw such residuals.* In graphs, or selected portions of graphs, where there are no directed inputs to a specific variable $\mathbf{v}_i$, the variable $\mathbf{v}_i$ is considered the same as its own residual $\mathbf{u}_i$, and the undirected turn represents relationships among $\mathbf{v}_i$ and other variables. However, directed arrows do end on $\mathbf{v}_i$; then v_i *is not the same as* u_i, and the undirected turn represents relationships among $\mathbf{u}_i$ and other points. In general graphic or algebraic terms, the interpretation of any structural arrow or turn always requires reference to the full graph or model.

COMPARATIVE REPRESENTATION

We now show how some popular models for moment structures may be completely specified under the simple RAM rules. We use both algebraic proofs and graphic presentation to further illustrate the general isomorphism between RAM algebra and graphics.

The models presented in Fig. 9.5 illustrate concepts revolving around the structure of relationships among manifest variables. Let us write a vector model for multiple linear regression as

$$\mathrm{y} = \mathrm{Bx} + \mathrm{e}, \tag{21}$$

where y is a $q \times 1$ vector of endogenous variables, x is a $p \times 1$ vector of exogeneous variables, and e is a $q \times 1$ vector of residuals. For simplicity, let us also assume that y and x are rescaled to have zero mean, so that B is a $q \times p$ matrix of raw regression coefficients for y regressed on x. Let us further assume model covariances

$$E\{\mathbf{x}\mathbf{x}'\} = \Sigma_{xx} \quad \text{and} \quad E\{\mathbf{e}\mathbf{e}'\} = \Sigma_{ee}, \tag{22}$$

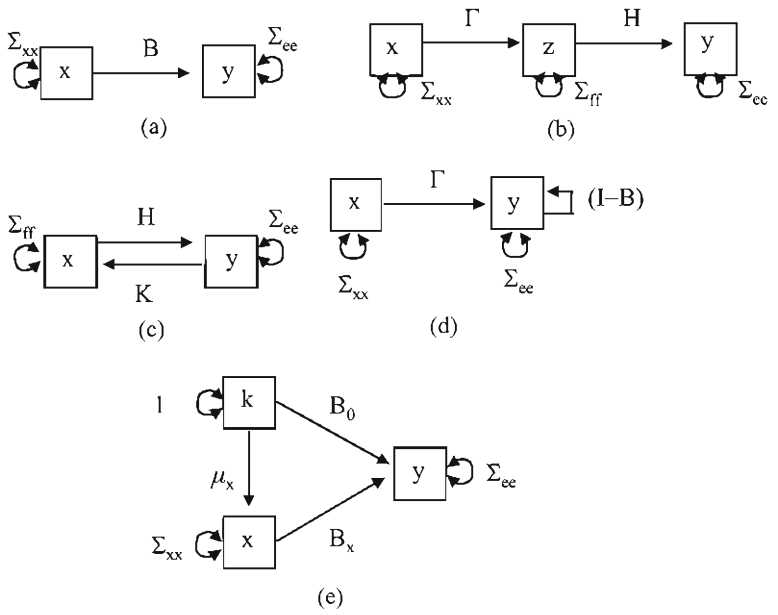

FIG. 9.5. Models of structure among manifest variables. (a) General linear regression, (b) chain or nested regression, (c) feedback or nonrecursive regression, (d) general econometric regression, (e) means and intercepts in a linear regression.

so that the covariance structure of $\mathbf{y}$ is obtained as

$$E\{\mathrm{y}\mathrm{y}'\} = \Sigma_{yy} = E\{(\mathrm{Bx} + \mathrm{e})(\mathrm{Bx} + \mathrm{e})'\} = \mathrm{B}E\{\mathrm{x}\mathrm{x}'\}\mathrm{B}'E\{\mathrm{e}\mathrm{e}'\} = \mathrm{B}\Sigma_{xx}\mathrm{B}' + \Sigma_{ee}. \tag{23}$$

In many familiar cases such as analysis of variance and covariance, we may wish to assume that Σ_{xx} is fully or partially represented in terms of fixed design parameters. However, in general, we can initially write the model without any additional restrictions.

This linear regression model may now be represented in RAM graphic terms as in Fig. 9.5a. Here it should be clear that the directed arrow represents the regression coefficients B, whereas the undirected turns represent covariance Σ_{xx} and residual covariances Σ_{ee} (not necessarily diagonal). *The latent variable e could be explicitly included in this diagram but it is not necessary* (e.g., Fig. 9.6c). By following the explicit graphic pathway rules, we can state the structure of the model-predicted associations among all variables in Equation 23 without further recourse to the matrix algebra. However, we can write this model in the compact form of *supervectors are with nonzero elements*

$$\begin{aligned} &j = t = p + q, \qquad \mathbf{v}' = [\mathrm{y}'\mathrm{x}'], \qquad \mathbf{u}' = [\mathrm{e}'\mathrm{x}'] \\ &\mathbf{F} = \begin{bmatrix} \mathrm{I} & \mathrm{O} \\ \mathrm{O} & \mathrm{I} \end{bmatrix}, \qquad \mathbf{A} = \begin{bmatrix} \mathrm{O} & \mathrm{B} \\ \mathrm{O} & \mathrm{O} \end{bmatrix}, \qquad \boldsymbol{\Omega} = \begin{bmatrix} \Sigma_{ee} & \text{sym.} \\ \mathrm{O} & \Sigma_{xx} \end{bmatrix}, \end{aligned} \tag{24}$$

so

$$\mathbf{E} = (\mathbf{I} - \mathbf{A})^{-1} = \begin{bmatrix} \mathrm{I} & \mathrm{B} \\ \mathrm{O} & \mathrm{I} \end{bmatrix}.$$

In general, we can simply state $\mathbf{F} = \mathbf{I}$ is assumed by $j = t$ (i.e., no explicit latent variables; see Equation 9), and the only nonzero subpartitions are those in row 1 and column 2 of $\mathbf{A}$ ($\mathbf{A}_{yx} = \mathrm{B}$) and in the ith row and ith column of $\boldsymbol{\Omega}$ ($\boldsymbol{\Omega}_{ii} = [\Sigma_{xx}\Sigma_{ee}]$). This yields the associations as

$$\begin{aligned} \boldsymbol{\Sigma} &= (\mathbf{I}_t - \mathbf{A})^{-1}\boldsymbol{\Omega}(\mathbf{I}_t - \mathbf{A})^{-1'} \\ &= \begin{bmatrix} \mathrm{I} & \mathrm{B} \\ \mathrm{O} & \mathrm{I} \end{bmatrix}\begin{bmatrix} \Sigma_{ee} & \mathrm{O} \\ \mathrm{O} & \Sigma_{\lambda\lambda} \end{bmatrix}\begin{bmatrix} \mathrm{I} & \mathrm{O} \\ \mathrm{B}' & \mathrm{I} \end{bmatrix} = \begin{bmatrix} \Sigma_{ee} + \mathrm{B}\Sigma_{xx}\mathrm{B}' & \text{sym.} \\ \Sigma_{xx}\mathrm{B}' & \Sigma_{\lambda\lambda} \end{bmatrix}. \end{aligned} \tag{25}$$

Obviously this result is equivalent to the covariance structure of Equations 22 and 23.

Let us now consider writing models for some traditional extensions of this general linear model. One model that stems from path-analytic research (Duncan, 1975) is a "nested" set of manifest variables, which may be written as

$$\mathrm{z} = \Gamma\mathrm{x} + \mathrm{e}, \qquad \mathrm{y} = \mathrm{H}\mathrm{z} + \mathrm{f}, \tag{26}$$

where we assume x, y, and z are manifest and e and f are residuals. We note that this model implies no direct structural regression between x and y, and for this reason z is termed a mediator and the model is termed a "chain" (Wold, 1960). This model is drawn in compact form as in Fig. 9.5b and written in the abbreviated RAM notation as

$$
\begin{aligned}
&j = t, \qquad \mathbf{v}' = [\mathrm{x}'\mathrm{z}'\mathrm{y}'], \qquad \mathbf{u}' = [\mathrm{x}'\mathrm{e}'\mathrm{f}'], \\
&\mathbf{F} = \begin{bmatrix} \mathrm{I} & \mathrm{O} & \mathrm{O} \\ \mathrm{O} & \mathrm{I} & \mathrm{O} \\ \mathrm{O} & \mathrm{O} & \mathrm{I} \end{bmatrix}, \qquad \mathbf{A} = \begin{bmatrix} \mathrm{O} & \mathrm{O} & \mathrm{O} \\ \Gamma & \mathrm{O} & \mathrm{O} \\ \mathrm{O} & \mathrm{H} & \mathrm{O} \end{bmatrix}, \\
&\Omega = \begin{bmatrix} \Sigma_{xx} & & \text{sym.} \\ \mathrm{O} & \Sigma_{ee} & \\ \mathrm{O} & \mathrm{O} & \Sigma_{ff} \end{bmatrix}, \qquad \mathbf{E} = (\mathbf{I} - \mathbf{A})^{-1} = \begin{bmatrix} \mathrm{I} & \mathrm{O} & \mathrm{O} \\ \Gamma & \mathrm{I} & \mathrm{O} \\ \Gamma\mathrm{H} & \mathrm{H} & \mathrm{I} \end{bmatrix}.
\end{aligned} \tag{27}
$$

where the parameter matrices are in the lower diagonal of $\mathbf{A}$ (and $\mathbf{E}$) due to the ordering of the variables in $\mathbf{v}$. This expression leads to the nonzero association submatrices written as

$$
\begin{aligned}
\Sigma_{xz} &= \Sigma_{xx}\Gamma, \\
\Sigma_{zz} &= \Sigma_{ee} + \Gamma\Sigma_{xx}\boldsymbol{\Gamma}', \\
\Sigma_{yx} &= H\Gamma\Sigma_{xx}, \\
\Sigma_{yz} &= H\Gamma\Sigma_{xx}\Gamma' + H\Sigma_{ee}, \\
\Sigma_{yy} &= H\Gamma\Sigma_{xx}\Gamma'H' + H\Sigma_{ee}H' + \Sigma_{ff} = H(\Gamma\Sigma_{xx}\Gamma' + \Sigma_{ee})H' + \Sigma_{ff}.
\end{aligned} \tag{28}
$$

Another model that has been studied extensively in path-analytic research is the feedback model (Heise, 1975), which may be written in terms of the nonrecursive equations

$$
y = Hz + f, \qquad z = Ky + s, \tag{29}
$$

where both y and z are manifest and f and s are residuals. This model is drawn as Fig. 9.5c and can be written as

$$
\begin{aligned}
&j = t, \qquad \mathbf{v}' = [\mathrm{z}'\mathrm{y}'], \qquad \mathbf{u} = [\mathrm{f}'\mathrm{s}'], \\
&\mathbf{F} = \begin{bmatrix} \mathrm{I} & \mathrm{O} \\ \mathrm{O} & \mathrm{I} \end{bmatrix}, \qquad \mathbf{A} = \begin{bmatrix} \mathrm{O} & \mathrm{K} \\ \mathrm{H} & \mathrm{O} \end{bmatrix}, \qquad \Omega = \begin{bmatrix} \Sigma_{ff} & \text{sym.} \\ \mathrm{O} & \Sigma_{ss} \end{bmatrix}, \quad \text{so} \\
&\mathbf{E} = (\mathbf{I} - \mathbf{A})^{-1} = \begin{bmatrix} \Xi & K\Xi \\ H\Xi & \Xi \end{bmatrix},
\end{aligned} \tag{30}
$$

where $\Xi = (I - HK)^{-1}$, so that nonzero submatrices yield a complementary covariance structure,

$$\begin{aligned} \Sigma_{zz} &= \Xi\Sigma_{ff}\Xi' + H\Xi\Sigma_{ss}\Xi'H' \\ \Sigma_{zy} &= K\Xi\Sigma_{ff}\Xi' + \Xi\Sigma_{ss}\Xi'H' \\ \Sigma_{yy} &= \Xi\Sigma_{ff}\Xi' + K\Xi\Sigma_{ff}\Xi'K'. \end{aligned} \tag{31}$$

It becomes obvious that the pathway termed Ξ, often considered as a feedback loop, is most crucial in this model. A complete decomposition of the separate pathways within this feedback requires the infinite sum of pathway products of ever-increasing length (see Equation 8 for a general representation). However, both algebraic and graphic forms are still equivalent; backward entry is possible only through Ξ and forward outcome only through Ξ' (Kenny, 1979, p. 100).

A general representation that permits this kind of feedback, as well as other structures, has been of particular interest to econometric research (Wold, 1960) in the form of nonrecursive equations, classically written as

$$By = \Gamma x + z \qquad \text{or} \qquad y = (I - B)y + \Gamma x + z, \tag{32}$$

where $(I - B)$ is now a $q \times q$ matrix of regression weights for y on itself. This expression is drawn as Fig. 9.5d and written as

$$\begin{aligned} &j = t = q + p, \qquad \mathbf{v}' = [y'x'], \qquad \mathbf{u}' = [z'x'], \\ &\mathbf{F} = \begin{bmatrix} I & O \\ O & I \end{bmatrix}, \qquad \mathbf{A} = \begin{bmatrix} (I - B) & O \\ \Gamma & O \end{bmatrix}, \qquad \mathbf{\Omega} = \begin{bmatrix} \Sigma_{zz} & \text{sym.} \\ O & \Sigma_{xx} \end{bmatrix}, \quad \text{so} \\ &\mathbf{E} = (\mathbf{I} - \mathbf{A})^{-1} = \begin{bmatrix} B^{-1} & O \\ B^{-1}\Gamma & I \end{bmatrix}, \end{aligned} \tag{33}$$

with the resulting expectations

$$\begin{aligned} \Sigma_{yy} &= B^{-1}\Sigma_{xx}B^{-1'} + B^{-1}\Gamma\Sigma_{zz}\Gamma'B^{-1'} = B^{-1}(\Sigma_{xx} + \Gamma\Sigma_{zz}\Gamma')B^{-1'}, \\ \Sigma_{yx} &= \Sigma_{zz}\Gamma'B^{-1'}. \end{aligned} \tag{34}$$

Once again, the feedback matrix $(I - B)$ among the **y** variables is seen as a crucially important part of the structural associations.

Let us finally consider a manifest variable model that includes both first and second moments for j variables with means and covariances. One simple approach is to write a linear regression model, such as Equation 16, with the explicit inclusion of means as

$$y = B_0 + B_x x + e, \tag{35}$$

where B_0 is a $p \times 1$ vector of regression coefficients representing the regression of y on the unit vector $k = 1$ (i.e., the mean intercepts). Let us also define

$$\begin{aligned} E\{x\} &= \mu_x, \qquad E\{e\} = 0, \qquad \text{so} \qquad E\{\mathrm{y}\} = \mathrm{B}_0 + \mathrm{B}_x\mu_x, \\ E\{k\} &= 1, \qquad E\{kk'\} = 1. \end{aligned} \tag{36}$$

To write this model in RAM notation, we require the explicit inclusion of a design structure where the last row, labeled 1 here, contains all unit values. Now we write the model

$$\begin{aligned} j &= t = p + q + 1, \qquad \mathbf{v}' = [\mathrm{y}'\mathrm{x}'\mathrm{k}'], \qquad \mathbf{u}' = [e'\mathrm{x}'\mathrm{k}'], \\ \mathbf{F} &= \begin{bmatrix} \mathrm{I} & \mathrm{O} & \mathrm{O} \\ \mathrm{O} & \mathrm{I} & \mathrm{O} \\ \mathrm{O} & \mathrm{O} & \mathrm{I} \end{bmatrix}, \qquad \mathbf{A} = \begin{bmatrix} \mathrm{O} & \mathrm{B}_x & \mathrm{B}_0 \\ \mathrm{O} & \mathrm{O} & \mathrm{M}_x \\ 0 & 0 & 0 \end{bmatrix}, \boldsymbol{\Omega} = \begin{bmatrix} \Sigma_{ee} & & \text{sym.} \\ \mathrm{O} & \Sigma_{xx} & \\ 0 & 0 & 1 \end{bmatrix} \end{aligned} \tag{37}$$

so

$$\mathbf{E} = (\mathbf{I} - \mathbf{A})^{-1} = \begin{bmatrix} \mathrm{I} & \mathrm{B}_x & \mathrm{B}_0 \\ \mathrm{O} & \mathrm{I} & (\mathrm{B}_0 + \mathrm{B}_x\mu_x) \\ 0 & 0 & 1 \end{bmatrix},$$

and the matrices of expectations are written in submatrices

$$\begin{aligned} \Sigma^*_{yy} &= \Sigma_{\mathrm{ee}} + \mathrm{B}_x\Sigma_{xx}\mathrm{B}'_x + (\mathrm{B}_x\mu_x + \mathrm{B}_0)(\mathrm{B}_x\mu_x + \mathrm{B}_0)', \\ \Sigma^*_{yx} &= \Sigma_{xx}\mathrm{B}'_x + \mu_x(\mathrm{B}_x\mu_x + \mathrm{B}_0)', \\ \Sigma^*_{xx} &= \Sigma_{xx} + \mu_x\mu'_x, \qquad \Sigma^*_{yk} = (\mathrm{B}_x\mu_x + \mathrm{B}_0), \\ \Sigma^*_{xk} &= \mu_x, \qquad \Sigma^*_{kk} = 1. \end{aligned} \tag{38}$$

As Fig. 9.5e shows, when the means are represented as regressions from a constant they follow all algebraic and graphic rules. A moment matrix that is augmented by a unit constant $\boldsymbol{\Sigma}^*$ of order $j^* = j + 1$ permits the complete separation of first and second moments. In general, all moment structure models may be written and interpreted in this way. Because unit constants are treated in this special fashion, highlighting these "defined" constants in the graphic (e.g., as a triangle) may be useful.

The models presented in Fig. 9.6 illustrate concepts based on the structure of *relations including latent variables*. One such model that has been of general interest in psychological research is the traditional model of common factor analysis (Thurstone, 1947), usually written as

$$\mathrm{y} = \Lambda\mathrm{f} + \mathrm{s}, \tag{39}$$

where y is a $p \times 1$ vector of observed variables, f is an $r \times 1$ vector of unobserved common factors, and s is a $p \times 1$ vector of unobserved unique factors (i.e., specific factors plus errors). Let us again assume that these variables have mean zero,

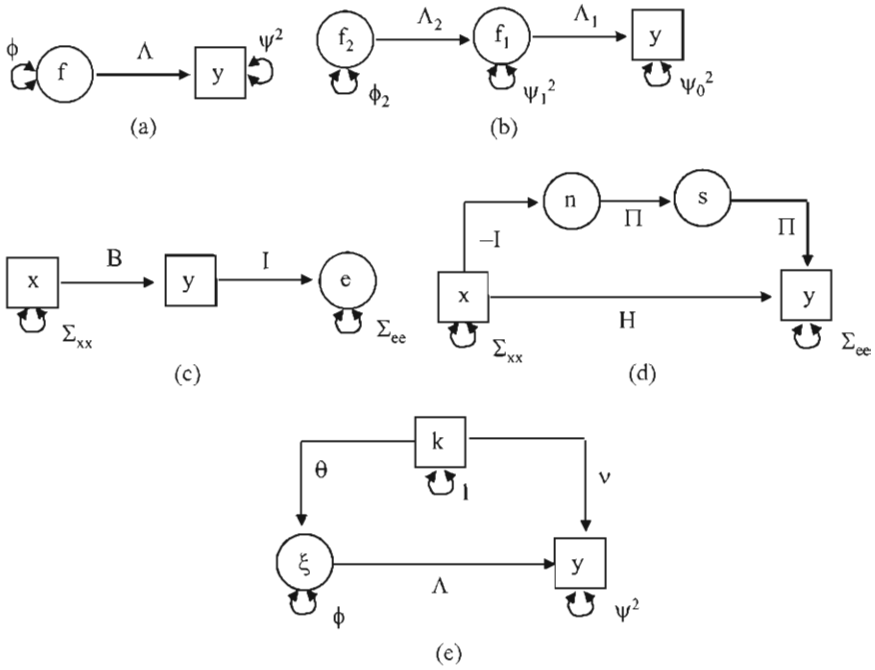

FIG. 9.6. Models of structure among both manifest and latent variables. (a) Traditional common factor model. (b) higher order factor model. (c) explicit inclusion of the residual in the regression. (d) linear constraints in the regression, (e) means and intercepts in a common factor model.

so that Λ is a $p \times r$ matrix of common factor loadings. Under these traditional assumptions, we may draw a general RAM model as in Fig. 9.6a and write

$$
\begin{aligned}
& j = p, \qquad k = r, \qquad \mathbf{v}' = [y'f'], \qquad \mathbf{u}' = [s'f'], \\
& \mathbf{F} = \begin{bmatrix} \mathrm{I} & \mathrm{O} \end{bmatrix}, \qquad \mathbf{A} = \begin{bmatrix} \mathrm{O} & \Lambda \\ \mathrm{O} & \mathrm{O} \end{bmatrix}, \qquad \boldsymbol{\Omega} = \begin{bmatrix} \Psi^2 & \text{sym.} \\ \mathrm{O} & \Phi \end{bmatrix}, \quad \text{so} \\
& \mathbf{E} = (\mathbf{I} - \mathbf{A})^{-1} = \begin{bmatrix} \mathrm{I} & \Lambda \\ \mathrm{O} & \mathrm{I} \end{bmatrix},
\end{aligned}
\tag{40}
$$

and we obtain the expectations

$$
\begin{aligned}
\Sigma_{yy} &= \Psi + \Lambda \Phi \Lambda', \\
\Sigma_{yf} &= \Phi \Lambda', \\
\Sigma_{ff} &= \Phi.
\end{aligned}
\tag{41}
$$

The overall associations include the traditional factor model as well as the factor structure associations. Now, however, due to the fact that we are explicitly including latent variables, we must also write the filter $\mathbf{F} = [I : 0]$ to subpartition $\boldsymbol{\Sigma}$ into $\boldsymbol{\Sigma}_{yy} = \mathbf{F}\boldsymbol{\Sigma}\mathbf{F}'$.

The basic concept involved in nested sets of latent variables forms the basis of hierarchical factor analysis models (Schmid & Leiman, 1957). We consider a model written as

$$\mathbf{y} = \boldsymbol{\Lambda}_1\mathbf{f}_1 + \mathbf{s}_1, \qquad \mathbf{f}_1 = \boldsymbol{\Lambda}_2\mathbf{f}_2 + \mathbf{s}_2, \tag{42}$$

where $\mathbf{f}_1$ is an $r \times 1$ vector of first-order latent common factors, $\mathbf{f}_2$ is a $q \times 1$ vector of second-order latent common factors, $\mathbf{s}_1$ and $\mathbf{s}_2$ are first- and second-order unique factors, respectively, and $\boldsymbol{\Lambda}_1$ and $\boldsymbol{\Lambda}_2$ are first- and second-order factor loading matrices, respectively. This model may be simply drawn as in Fig. 9.6b and written as

$$\begin{gathered} j = p, \qquad j = r + q, \qquad \mathbf{v}' = [\mathrm{y}'\mathrm{f}_1'\mathrm{f}_2'], \qquad \mathbf{u}' = [\mathrm{s}_1'\mathrm{s}_2'\mathrm{f}_2'], \\ \mathbf{F} = \begin{bmatrix} \mathrm{I} & \mathrm{O} & \mathrm{O} \end{bmatrix}, \qquad \mathbf{A} = \begin{bmatrix} \mathrm{O} & \Lambda_1 & \mathrm{O} \\ \mathrm{O} & \mathrm{O} & \Lambda_2 \\ \mathrm{O} & \mathrm{O} & \mathrm{O} \end{bmatrix}, \qquad \boldsymbol{\Omega} = \begin{bmatrix} \Psi_1^2 & & \text{sym.} \\ \mathrm{O} & \Psi_2^2 & \\ \mathrm{O} & \mathrm{O} & \Phi_2 \end{bmatrix} \end{gathered} \tag{43}$$

so

$$\mathbf{E} = (\mathbf{I} - \mathbf{A})^{-1} = \begin{bmatrix} I & \Lambda_1 & \Lambda_1\Lambda_2 \\ O & I & \Lambda_2 \\ O & O & I \end{bmatrix},$$

with the resulting expectations

$$\begin{aligned} \Sigma_{yy} &= \Psi_1^2 + \Lambda_1\Psi_2^2\Lambda_1' + \Lambda_1\Lambda_2\Phi_2\Lambda_2'\Lambda_1' \\ &= \Psi_1^2 + \Lambda_1\left(\Psi_2^2 + \Lambda_2\Phi_2\Lambda_2'\right)\Lambda_1' \\ \Sigma_{y,f1} &= \Psi_2^2\Lambda_1' + \Lambda_2\Phi_2\Lambda_2\Lambda_1', \\ \Sigma_{f1,f1} &= \Psi_2^2 + \Lambda_2\Phi_2\Lambda_2', \\ \Sigma_{y,f2} &= \Phi_2\Lambda_2'\Lambda_1', \\ \Sigma_{f1,f2} &= \Phi_2\Lambda_2', \\ \Sigma_{f2,f2} &= \Phi_2. \end{aligned} \tag{44}$$

These associations represent a nested set of factor analysis structures for both first and second orders. Additional hierarchical structures (e.g., Bentler, 1976) merely require repeated sets of latent variables nested in this fashion and do not require special programming (Weeks, 1978).

To illustrate some further concepts in RAM theory, let us once again consider a model that is equivalent to the linear regression model of Fig. 9.5a, but that *now*

explicitly includes the residual **e** *as a special latent variable.* We may write this model as

$$
\begin{aligned}
&j = p + q, \qquad k = q, \qquad \mathbf{v}' = [\mathrm{y}'\mathrm{x}'\mathrm{e}'], \qquad \mathbf{u}' = [\mathrm{y}^{*\prime}\mathrm{x}'\mathrm{e}'], \\
&\mathbf{F} = \begin{bmatrix} \mathrm{I} & \mathrm{I} & \mathrm{O} \end{bmatrix}, \qquad \mathbf{A} = \begin{bmatrix} \mathrm{O} & \mathrm{B} & \mathrm{O} \\ \mathrm{O} & \mathrm{O} & \mathrm{I} \\ \mathrm{O} & \mathrm{O} & \mathrm{O} \end{bmatrix}, \qquad \boldsymbol{\Omega} = \begin{bmatrix} \mathrm{O} & & \text{sym.} \\ \mathrm{O} & \Sigma_{xx} & \\ \mathrm{O} & \mathrm{O} & \Sigma_{ee} \end{bmatrix},
\end{aligned} \tag{45}
$$

so there is a zero main diagonal element to signify the lack of any $\boldsymbol{\Omega}$ parameter for the **u** associated with this y variable. Here we obtain

$$
\begin{aligned}
\Sigma_{yy} &= \mathrm{B}\Sigma_{xx}\mathrm{B}' + \Sigma_{ee}, \\
\Sigma_{yx} &= \Sigma_{xx}\mathrm{B}', \qquad \Sigma_{ye} = \mathrm{I}\Sigma_{ee}\mathrm{I}'.
\end{aligned} \tag{46}
$$

Now, by a comparison of $\boldsymbol{\Sigma}$ in Equations 46 with $\boldsymbol{\Sigma}$ from the original linear regression model of Equation 25, we can see that this new model generates associations that are simply redundant with parameters already available in the model (i.e., the last row of $\boldsymbol{\Sigma}$ filled with Σ_{ee}). Also, because **e** is now explicitly written as a latent variable, we must now write a subpartitioned **F** to filter out these redundancies. In this algebraic sense, then, it is unnecessary to explicitly include latent residuals with unit regressions. The model of Fig. 9.6c *is a consistent expression and may be useful for didactic purposes* (i.e., for clarity of $E\{\mathrm{xe}'\} = 0$).

At this point, however, we can also consider some nonredundant uses of the explicit inclusion of latent residuals. For example, the linear regression model drawn in Fig. 9.6c may now be altered to a model where we add a diagonal matrix T of regressions for **y** on **e** by writing

$$
\begin{aligned}
&j = p + q, \quad k = q, \qquad \mathbf{v}' = [\mathrm{y}'\mathrm{x}'\mathrm{e}'], \qquad \mathbf{u}' = [\mathrm{y}^{*\prime}\mathrm{x}'\mathrm{e}'], \\
&\mathbf{F} = \begin{bmatrix} \mathrm{I} & \mathrm{I} & \mathrm{O} \end{bmatrix}, \quad \mathbf{A} = \begin{bmatrix} \mathrm{O} & \mathrm{B} & \mathrm{O} \\ \mathrm{O} & \mathrm{O} & \mathrm{T} \\ \mathrm{O} & \mathrm{O} & \mathrm{O} \end{bmatrix}, \qquad \boldsymbol{\Omega} = \begin{bmatrix} \mathrm{O} & & \text{sym.} \\ \mathrm{O} & \Sigma_{xx} & \\ \mathrm{O} & \mathrm{O} & \mathrm{I} \end{bmatrix},
\end{aligned} \tag{47}
$$

so by this substitution of T (and the removal of $\boldsymbol{\Sigma}_{\mathbf{ee}}$) we obtain

$$
\begin{aligned}
\Sigma_{yy} &= \mathrm{B}\Sigma_{xx}\mathrm{B}' + \mathrm{TT}', \\
\Sigma_{yx} &= \Sigma_{xx}\mathrm{B}', \qquad \Sigma_{ye} = \mathrm{IT}', \quad \Sigma_{ee} = \mathrm{TT}'.
\end{aligned} \tag{48}
$$

Although this model has similar properties to the previous one, it exhibits one important distinction; the residual variance $\Sigma_{ee} = \mathrm{TT}'$ elements are now constrained to be nonnegative because the square roots (deviations) T are structured instead of

Σ_{ee}. This model, then, places a nonnegative constraint on the variance term and may be useful in practice (van Driel, 1978).

Several other forms of general linear *constraints* may be modeled in precisely this fashion. Let us consider a linear regression model with constraints on the B coefficients drawn as in Fig. 9.6d and written as

$$\begin{gathered} j = p + q, \qquad k = p + p, \qquad \mathbf{v}' = [y'x'n's'], \qquad \mathbf{u}' = [z'x'n's'], \\ \mathbf{F} = \begin{bmatrix} I & O & O & O \\ O & I & O & O \end{bmatrix}, \qquad \mathbf{A} = \begin{bmatrix} O & H & -I & O \\ O & O & O & O \\ O & \Pi & O & O \\ O & O & \Pi & O \end{bmatrix}, \\ \boldsymbol{\Omega} = \begin{bmatrix} \Sigma_{ee} & & & \text{sym.} \\ O & \Sigma_{xx} & & \\ O & O & O & \\ O & O & O & O \end{bmatrix}, \end{gathered} \tag{49}$$

where s and n are latent vectors with expectation zero and the regression matrix Π is diagonal, so that $\Sigma_{yx} = (H - \Pi\Pi) = B^*$ and

$$\begin{aligned} \Sigma_{yy} &= \Sigma_{ee} + B^* \Sigma_{xx} B^{*\prime}, \\ \Sigma_{yx} &= \Sigma_{xx} B^{*\prime}. \end{aligned} \tag{50}$$

The key result here is that this model yields structural associations that are equivalent to the linear regression model given in Equation 25 except for the fact that we have restructured B with $B^* = (H - \Pi\Pi)$. Now, because Π is structured in this way, the value $\Pi\Pi \geq 0$ and this model *constrains the value* $B^* \leq H$, *where* H *may be any fixed at a known value or the value of another parameter*. The fact that neither n nor s has any has associated expectation, and thus is "not variable," poses no problems for the RAM representation. In such cases we will just label n and s as *nodes* (following Horn & McArdle, 1980). The utility of the explicit inclusion of such nodes is that they clearly permit the direct representation of this variety of model constraint, as well as other more complex combinations that follow directly. On the other hand, we also recognize that not all nonlinear model constraints may be accomplished in this simple fashion (McDonald, 1980).

Given these basic relations, we can now consider more general linear structural models. Following Sörbom (1974; also Jöreskog & Sörbom, 1980), we may wish to write the factor analysis model with means as

$$y = \nu + \Lambda\eta + \epsilon, \qquad E\{\eta\} = \theta, \qquad E\{\epsilon\} = 0, \qquad \text{so} \quad E\{y\} = \nu + \Lambda\theta, \tag{51}$$

where ν represents manifest variable mean intercepts and θ represents the means of common factors η. This model, presented in Fig. 9.6e, may be represented

in a fashion equivalent to the mean regression model given in Fig. 9.5e. That is, we create a constant unit vector k (now ordered as the first variable) and we write

$$j = 1 + p, \qquad k = r, \qquad \mathbf{v}' = [\mathrm{k}'\mathrm{y}'\eta'], \qquad \mathbf{u}' = [\mathrm{k}'\epsilon'\eta'],$$
$$\mathbf{F} = \begin{bmatrix} 1 & O & O \\ 0 & I & O \end{bmatrix}, \qquad \mathbf{A} = \begin{bmatrix} 0 & O & O \\ \nu & O & \Lambda \\ \theta & O & O \end{bmatrix}, \qquad \Omega = \begin{bmatrix} 1 & & \text{sym.} \\ O & \Psi^2 & \\ O & O & \Phi \end{bmatrix}, \tag{52}$$

so

$$\mathbf{E} = (\mathbf{I} - \mathbf{A})^{-1} = \begin{bmatrix} I & O & O \\ (\nu + \Lambda\theta) & I & O \\ \theta & O & I \end{bmatrix},$$

where the constant k is not filtered, so

$$\Sigma^*_{kk} = 1, \qquad \Sigma_{ky} = (\nu + \Lambda\theta)$$
$$\Sigma^*_{yy} = \Psi + \Lambda\Phi\Lambda' + (\nu + \Lambda\theta)(\nu + \Lambda\theta)', \tag{53}$$
$$\Sigma^*_{k\eta} = \theta, \qquad \Sigma^*_{y\eta} = \Phi\Lambda + \theta(\nu + \Lambda\theta)', \qquad \Sigma^*_{\eta\eta} = \Phi + \theta\theta'.$$

In general, any regression from the constant vector is scaled in terms of first moments for either manifest or latent variables, or both simultaneously (Sörbom, 1978).

INTERRELATIONSHIPS AMONG GENERAL PURPOSE MODELS

Let us now turn to a consideration of some important principles involving the *interrelationships among the general purpose models* presented in Fig. 9.7. The first model to be considered is the recent, but very popular modeling foundation presented by Jöreskog (1973b) as the linear structural relations (LISREL) model (also see Jöreskog & Sörbom, 1980; Wiley, 1973). This model, when drawn as in Fig. 9.7a, can clearly be seen as a combination of the econometric nonrecursive model given in Fig. 9.5c and the psychometric factor analysis model given in Fig. 9.6a. This model is usually written as

$$\mathrm{x} = \Lambda_x\xi + \delta, \qquad \mathrm{y} = \Lambda_y\eta + \epsilon, \qquad \text{and} \qquad \mathrm{B}\eta = \Gamma\xi + \zeta, \tag{54}$$

where x ($\mathrm{p} \times 1$) and y ($\mathrm{q} \times 1$) represent observed variables, ξ ($\mathrm{n} \times 1$) and η ($\mathrm{m} \times 1$) represent unobserved common factors, δ and ε represent unique factors, and ζ represents a structural residual variable. Here, Λ_x and Λ_y are factor pattern matrices, and B and Γ are structural regression coefficient matrices. As

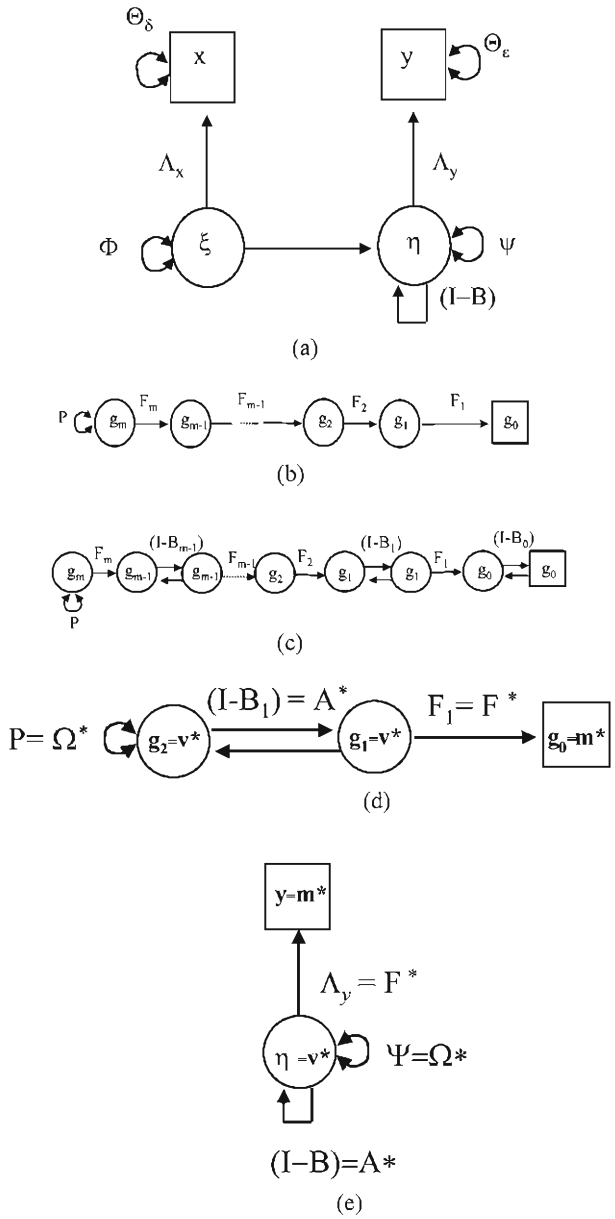

FIG. 9.7. General models represented in reticular action meta-model (RAM) notation. (a) The standard linear structural relations (LISREL) model, (b) the standard covariance structure analysis (COSAN) model, (c) the complete COSAN model, (d) the RAM–COSAN interface, (d) the RAM–LISREL interface.

McArdle (1978) showed, we can write this model in RAM notation as

$$j = p + q, \qquad k = n + m, \qquad \mathbf{v}' = [\mathrm{y}'\mathrm{x}'\eta'\xi'], \qquad \mathbf{u}' = [\epsilon\delta'\zeta'\xi'],$$

$$\mathbf{F} = \begin{bmatrix} \mathrm{I} & \mathrm{O} & \mathrm{O} & \mathrm{O} \\ \mathrm{O} & \mathrm{O} & \mathrm{O} & \mathrm{O} \end{bmatrix}, \qquad \mathbf{A} = \begin{bmatrix} \mathrm{O} & \mathrm{O} & \Lambda_y & \mathrm{O} \\ \mathrm{O} & \mathrm{O} & \mathrm{O} & \Lambda_x \\ \mathrm{O} & \mathrm{O} & (\mathrm{I} - \mathrm{B}) & \Gamma \\ \mathrm{O} & \mathrm{O} & \mathrm{O} & \mathrm{O} \end{bmatrix}, \tag{55}$$

$$\mathbf{\Omega} = \begin{bmatrix} \Theta_{\varepsilon\varepsilon} & & & \text{sym.} \\ \mathrm{O} & \Theta_{\delta\delta} & & \\ \mathrm{O} & \mathrm{O} & \Psi & \\ \mathrm{O} & \mathrm{O} & \mathrm{O} & \Phi \end{bmatrix},$$

so

$$\mathbf{E} = (\mathbf{I} - \mathbf{A})^{-1} = \begin{bmatrix} \mathrm{I} & \mathrm{O} & \Lambda_y \mathrm{B}^{-1} & \Lambda_y \mathrm{B}^{-1}\Gamma \\ \mathrm{O} & \mathrm{I} & \mathrm{O} & \Lambda_x \\ \mathrm{O} & \mathrm{O} & \mathrm{B}^{-1} & \mathrm{B}^{-1}\Gamma \\ \mathrm{O} & \mathrm{O} & \mathrm{O} & \mathrm{O} \end{bmatrix},$$

and so

$$\begin{aligned} \Sigma_{yy} &= \Lambda_y \Sigma_{\eta\eta} \Lambda_y' + \Theta_\varepsilon, \\ \Sigma_{yx} &= \Lambda_x \mathrm{B}^{-1} \Gamma \Phi \Lambda_y', \\ \Sigma_{xx} &= \Lambda_x \Phi \Lambda_x' + \Theta_\delta', \\ \Sigma_{y\eta} &= \Sigma_{\eta\eta} \Lambda_y', \\ \Sigma_{x\eta} &= \Sigma_{\eta\eta} \Phi \Lambda_x' \\ \Sigma_{\eta\eta} &= \mathrm{B}^{-1} \Gamma \Phi \Gamma' \mathrm{B}^{-1'} + \mathrm{B}^{-1} \Psi \mathrm{B}^{-1'} = \mathrm{B}^{-1} (\Gamma \Phi \Gamma' + \Psi) \mathrm{B}^{-1'}, \\ \Sigma_{y\xi} &= \Phi \Sigma_{\eta\eta} \Lambda_y', \\ \Sigma_{x\xi} &= \Phi \Lambda_x', \qquad \Sigma_{\eta\xi} = \Phi \mathrm{B}^{-1} \Gamma, \qquad \Sigma_{\xi\xi} = \Phi. \end{aligned} \tag{56}$$

This moment structure is equivalent to the manifest variable covariance structure proposed by Jöreskog (1973b, p. 87) and Wiley (1973, p. 80), and also includes all other latent and manifest variable associations. These matrices also demonstrate models that LISREL is not directly designed to fit (i.e., the zero locations in matrixes **A** and $\mathbf{\Omega}$).

We may also wish to develop the association matrix for multiple groups (as in Jöreskog, 1971; Sörbom, 1974, 1978) using a similar device. In the case of equal sample sizes we could simply create a matrix $\mathbf{\Sigma}^*$ that is appropriately subpartitioned into main diagonal block $\mathbf{\Sigma}_1, \mathbf{\Sigma}_2, \ldots, \mathbf{\Sigma}_g$, with zero off-diagonal blocks denoting independence between groups, and with all matrices scaled by the appropriate sample sizes. This supermatrix could then be analyzed under alternative structural representations of **A** and $\mathbf{\Omega}$ with appropriate constraints. It follows that Sörbom's (1978) inclusion of latent variable means combined within a multiple

group LISREL model may be accomplished by including a constant variable, as in Fig. 9.6e.

Another general model, first proposed by McDonald (1978; see also McDonald, 1980; McDonald & Swaminathan, 1972), is the covariance structure analysis (COSAN) model, which may be described as

$$\mathrm{g}_0 = \left(\prod_{j=1}^{m} F_j \right) \mathrm{g}_m, \tag{57}$$

where g_0 is an $h_0 \times 1$ vector of manifest variables, g_m is an $h_m \times 1$ vector of latent variables, and F_j is an $h_{j-1} \times h_j$ matrix of coefficients. By also writing the $h_m \times h_m$ symmetric matrix

$$E\{\mathrm{g}_m \, \mathrm{g}_m'\} = \mathrm{P}, \tag{58}$$

we may describe the manifest variable moment structure of COSAN as

$$\Sigma = E\{\mathrm{g}_0, \mathrm{g}_0'\} = \mathrm{F}_1\mathrm{F}_2 \cdots \mathrm{F}_m\mathrm{P}\mathrm{F}_m' \cdots \mathrm{F}_2'\mathrm{F}_1' = \left(\prod_{j=1}^{m} \mathrm{F}_j \right) \mathrm{P} \left(\prod_{j=1}^{m} \mathrm{F}_j \right)'. \tag{59}$$

When described in this direct hierarchical fashion, COSAN may be written in RAM algebra as

$$\begin{aligned} j &= h_0, \qquad k = \sum_{j=1,m} h_j, \qquad \mathbf{v}' = [\mathrm{g}_0'\mathrm{g}_1'\mathrm{g}_2' \cdots \mathrm{g}_m'], \\ \mathbf{u}' &= [\mathrm{g}_0^{*\prime}\mathrm{g}_1^{*\prime}\mathrm{g}_2^{*\prime} \cdots \mathrm{g}_m{}'], \qquad \mathbf{A}_{(j,j+1)} = \mathrm{F}_j, \qquad \Omega_{ii} = [0, 0, 0, \ldots, \mathrm{P}]. \end{aligned} \tag{60}$$

Figure 9.7b graphically shows that these general equations represent an mth-level latent hierarchy with $m - 1$ latent nodes.

This interrelationship would be complete except for the fact that COSAN also permits the representation of any model where

$$\mathrm{F}_j = \mathrm{B}^{-1} \qquad \text{and/or} \qquad \mathrm{P} = \mathrm{Q}^{-1}, \tag{61}$$

where B and/or Q may be directly patterned (Bentler & Weeks, 1979, p. 176). This extension of COSAN to include patterned inverses at any hierarchical order provides for much more complex models, such as LISREL (McDonald & Swaminathan, 1972). However, McArdle and McDonald (1984) showed that it is also possible to represent any COSAN model with patterned F_j inverses as a RAM model where

$$\begin{aligned} \mathrm{A}_{(j,j+1)} &= \mathrm{L}_j \quad \text{for } j = 1, \ldots, m \quad \text{and} \\ \mathrm{A}_{(i,i)} &= \mathrm{I} - \mathrm{B}_i \quad \text{for } i = 0, \ldots, m - 1. \end{aligned} \tag{62}$$

That is, patterned matrices such as $\mathbf{L}_i$ in Equations 62 should appear on the

superdiagonal of **A**, and patterned inverses B_i should appear on the main diagonal of **A**. This mixed-sequence arrangement can now yield any COSAN model by appropriately setting particular L_j or B_i equal to an identity matrix. The general form of this RAM–COSAN interchange is drawn as a path diagram in Fig. 9.7c.

As a specific example of this interchange, let us write a second-order COSAN model as

$$\Sigma = L_1 B_1^{-1'} P B_1^{-1'} L_1' \tag{63}$$

by initially writing a RAM model where

$$\mathbf{A}_{12} = L_1, \qquad \mathbf{A}_{11} = I - B_0, \qquad \text{and} \qquad \mathbf{A}_{22} = B_1, \tag{64}$$

so that by setting $B_0 = I$, we obtain $\Sigma = \mathbf{A}_{11}$ equivalent to Equation 63. However, as McArdle (1978) showed, now it is possible to write a COSAN model where we set

$$\begin{aligned} &m = 2, \qquad h_0 = j^*, \qquad h_1 = h_2 = t^*, \qquad g_0 = \mathbf{m}^*, \qquad g_1 = g_2 = \mathbf{v}^*, \\ &\mathbf{F}_1 = \mathbf{F}^*, \qquad \mathbf{F}_2 = B_2^{-1} = (I - \mathbf{A}^*)^{-1}, \qquad \text{and} \qquad P = \Omega^*, \end{aligned} \tag{65}$$

where asterisks are used to denote RAM vectors and matrices, and the substitution of Equations 65 into the COSAN result in Equation 59 yields the original RAM equation 13. This specific COSAN model is depicted in Fig. 9.7d.

This result naturally leads us to pursue the interrelationships between RAM and COSAN and the LISREL foundation. Specifically, we now know that we may write nonzero LISREL variables and parameter matrices as

$$\begin{aligned} &q = j^*, \qquad m = t^*, \qquad y = \mathbf{m}^*, \qquad \eta = \mathbf{v}^*, \qquad \zeta = \mathbf{u}^*, \\ &\Lambda_y = \mathbf{F}^*, \qquad B = (I - \mathbf{A}^*), \qquad \Psi = \Omega^*, \end{aligned} \tag{66}$$

where asterisks are used to denote RAM vectors and matrices and the substitution of Equations 66 into the LISREL Equations 54–58 also yields the RAM model of Equation 13. This LISREL model of RAM is presented in Fig. 9.7e. More details of the proof of this RAM–COSAN–LISREL interchange are presented by McArdle and McDonald (1984) and will be discussed in the last section.

COMPUTATIONAL REPRESENTATION

Due to the previous proof of the RAM–COSAN–LISREL interchange, theory concerning estimation, identification, and asymptotic testing of hypotheses in RAM can be taken directly from McDonald's (1978) COSAN or Jöreskog and Sörbom's (1978) LISREL. This is critical because, as McDonald (1978) said, "a convenient

simplification of the consequent mathematics . . . would be illusory if it did not carry over into the arithmetic of the estimation procedures" (p. 70). Technical results are presented here to illustrate how such simplifications are made possible through the use of RAM theory.

Directly paraphrasing McDonald (1978), let us say that the moment structure $\boldsymbol{\Sigma}$ is a prescribed function f of p mathematically independent parameters in $\mathbf{A}$ and $\boldsymbol{\Omega}$. Most specifically, we can say that the parameter estimates for $\mathbf{A}$ and $\boldsymbol{\Omega}$ are found among the solutions of the first-order partial derivatives for any RAM model listed as

$$\begin{aligned} \partial f(\mathbf{S}, \boldsymbol{\Sigma})/\partial \mathbf{A} &= \mathbf{I} - 2Vec\,([\mathbf{FE}]'\mathbf{G}[\mathbf{FE}\boldsymbol{\Omega}\mathbf{E}]) = 0, \\ \partial f(\mathbf{S}, \boldsymbol{\Sigma})/\partial \boldsymbol{\Omega} &= Vec\,([\mathbf{FE}\boldsymbol{\Omega}]'\mathbf{G}[\mathbf{FE}\boldsymbol{\Omega}]) = 0, \end{aligned} \tag{67}$$

where $\mathbf{S}$ is the sample moment matrix from a sample size N, $f(\mathbf{S}, \boldsymbol{\Sigma})$ is any loss function, $\mathbf{E} = (\mathbf{I} - \mathbf{A})^{-1}$ in as before, *Vec* $(\cdot)$ represents a matrix operator that lexically realigns all matrix elements into a vector, and $\mathbf{G} = \partial f/\partial \boldsymbol{\Sigma}$ is based on the explicit choice of a loss function. Note the $\partial \boldsymbol{\Sigma}/\partial \mathbf{A}$ and $\partial \boldsymbol{\Sigma}/\partial \boldsymbol{\Omega}$ may be easily written in the same fashion.

From a practical standpoint we can now develop a general algorithm for solving these equations merely by using the general techniques of nonlinear optimization. That is, following McDonald and Swaminathan (1972), we can write a general quadratic form of a Newton-based optimization scheme for the sth iterative step in a direct RAM solution as

$$\mathrm{P}(\mathbf{A},\ \boldsymbol{\Omega})^{s+1} = \mathrm{P}(\mathbf{A},\ \boldsymbol{\Omega})^{s} - d^{s}(\mathbf{M}^{-1})\mathbf{G}^{s}, \tag{68}$$

where $\mathrm{P}(\mathbf{A},\ \boldsymbol{\Omega})^{k}$ represents a vector of estimated values for $\mathbf{A}$ and $\boldsymbol{\Omega}$ parameters estimated at the kth step; the changes at each iteration are defined by a scaling constant d^{s} whose value is usually chosen in the interval (0, 1) to prevent descent divergence; $\mathbf{M}$ is a square symmetric positive-definite matrix; and $\mathbf{G}$ is the gradient vector. It is well known that the choice of correction elements in $\mathbf{M}$ may be defined in many ways to yield differential performance. For example, setting $\mathbf{M} = \mathbf{H}$ yields the classic Newton–Raphson procedure, which is known to behave quite well in the vicinity of the true solution. It is now a simple matter to write the second-order partial derivatives for RAM following McDonald (1978, pp. 68–69) as

$$\begin{aligned} \partial^2 f/\partial \mathbf{A}^2 &= ([\mathbf{FE}]' \otimes [\mathbf{FE}\boldsymbol{\Omega}\mathbf{E}]')\mathbf{JHJ}([\mathbf{FE}] \otimes [\mathbf{FE}\boldsymbol{\Omega}\mathbf{E}]) \\ &\quad + 2([\mathbf{FE}]'\mathbf{G}[\mathbf{FE}] \otimes [\mathbf{E}\boldsymbol{\Omega}\mathbf{E}'] \\ &\quad + \{\mathbf{E}' \otimes [\mathbf{FE}\boldsymbol{\Omega}\mathbf{E}']'\mathbf{G}[\mathbf{FE}] + [\mathbf{FE}]'\mathbf{G}[\mathbf{FE}\boldsymbol{\Omega}\mathbf{E}'] \otimes \mathbf{E}\}\mathbf{J}\} \\ \partial^2 f/\partial \boldsymbol{\Omega}^2 &= ([\mathbf{FE}\boldsymbol{\Omega}]' \otimes [\mathbf{FE}\boldsymbol{\Omega}]')\mathbf{H}([\mathbf{FEU}] \otimes [\mathbf{FEU}]), \\ \partial^2 f/\partial \boldsymbol{\Omega}\partial \mathbf{A} &= -\{([\mathbf{FE}\boldsymbol{\Omega}]' \otimes [\mathbf{FE}\boldsymbol{\Omega}]')\mathbf{HJ}([\mathbf{FE}] \otimes [\mathbf{FE}\boldsymbol{\Omega}\mathbf{E}']) \\ &\quad + 2[\mathbf{FE}\boldsymbol{\Omega}]'\mathbf{G}[\mathbf{FE}] \otimes [\mathbf{E}\boldsymbol{\Omega}]', \end{aligned} \tag{69}$$

where $\mathbf{H} = \partial^2 f / \partial \Sigma^2$ is also based on the choice of a loss function, $\otimes$ denotes the right-hand Kronecker product, and $\mathbf{J}$ is a permutation matrix of prescribed unity and zero values. Note that $\partial^2 \Sigma / \partial \mathbf{A}^2$, $\partial^2 \Sigma / \partial \Omega^2$, and $\partial^2 \Sigma / \partial \mathbf{A}\, \partial \Omega$ may be written in the same fashion. In any case, it should be obvious that *these equations are simplifications of the general treatments* of McDonald (1978) and Jöreskog and Sörbom (1978), yet they can be used to yield identical results. Alternative algorithmic solutions that have been shown to exhibit good comparative performance and do not require these second order derivatives include the Gauss–Newton method (Dennis & Moré, 1978; Lee, 1977) and the Davidon–Fletcher–Powell method (Jöreskog & Sorbom, 1978).

The practical efficiency of a general algorithm is of major concern. This turns out to be a critical issue in RAM theory due to the fact that a general solution for the RAM model will require the repetitive matrix multiplication and inversion of many large and sparse matrices. For this reason, the use of specialized routines for *sparse matrix operations* (Gustavson, 1978) and matrix storage (Duff & Reid, 1979) will no doubt be a required feature of any generally practical algorithm. In all cases, the calculation of $\mathbf{A}$ should be done noting the symmetry of Ω, and, due to the unusually simple pattern inherent in $\mathbf{F}$, the calculation of Σ from $\mathbf{A}$ should be done without direct multiplication by $\mathbf{F}$.

The use of the current sparse matrix algorithm for the repetitive calculation of $\mathbf{E} = (\mathbf{I} - \mathbf{A})^{-1}$ will, for the most part, be an attempt to translate the expensive inverse calculations into a series of much more efficient sparse matrix multiplications. This is usually done by searching for specific patterns (e.g., diagonal block) in the original sparse matrix (Gustavson, 1978). It follows that the more efficiently such patterns are found, the more efficient the overall algorithm will perform. In a large variety of problems it will be possible to simply reorder the vector $\mathbf{v}$ so that $\mathbf{A}$ is upper or lower triangular, and the inverse may be calculated on a single pass of a sweep operation (Jennrich, 1977).

A practical solution to this problem is now possible using RAM theory; the computational inversion of $(\mathbf{I} - \mathbf{A})$ may capitalize on the fundamental relationships among the general models previously described to gain some efficiency in programming. That is, we can now create an algorithm that will (a) efficiently describe the patterns in the RAM matrix $\mathbf{A}$, (b) directly translate these patterns into a series of smaller sparse matrix multiplications using COSAN calculations, and (c) directly translate final estimates back to the RAM notation. The most crucial step in this new algorithm is the direct translation of any arbitrary RAM $\mathbf{A}$ and Ω matrices into an equivalent series of COSAN F_j and P matrices. Equation 62 illustrates some of the basic principles in the interface between RAM and COSAN, but it does not provide a direct answer to the reverse problem of going from RAM parameters to COSAN parameters. The main steps in a translation from a RAM model to a COSAN model are direct, merely requiring the repetitive search through $\mathbf{A}$ and Ω for specific patterns. However, the complete details of this algorithmic pattern search are too lengthy to detail here (McArdle & Horn, 1981).

To illustrate the possible efficiency of this RAM–COSAN interchange, let us now reconsider the multiple linear regression model of RAM Equations 24 and 25 expressed as a COSAN model in abbreviated notation as

$$\begin{aligned} m &= 1, \qquad g_0' = [y'x'], \qquad g_1' = [e'x'], \\ F_1 &= L_1 = \begin{bmatrix} I & B \\ O & I \end{bmatrix}, \qquad P = \begin{bmatrix} \Sigma_{ee} & \text{sym.} \\ O & \Sigma_{xx} \end{bmatrix}. \end{aligned} \tag{70}$$

This model, as may be verified by substitution in COSAN Equation 59, yields a moment structure that is equivalent to Equation 24. This translation to COSAN requires several elemental numerical operations based on the size of y and x, but the COSAN formulation now does not require the repetitive inversion of any model matrix. Thus, we have now effectively replaced the iterative calculation of $\mathbf{E} = (\mathbf{I} - \mathbf{A})^{-1}$ with a single multiplication problem of an equivalent size that is more easily handled by sparse matrix routines. This exact result applies for the factor analysis model written as Equations 39–41 and the model with the explicit inclusion of errors written as Equations 45 and 46.

As another example, we can translate the linear nested equation path model of RAM Equations 27 and 28 into a COSAN model of

$$\begin{aligned} m &= 2, \qquad g_0' = [x'y'z'], \qquad g_1' = [x'y'f'], \qquad g_2' = [x'q'f'], \\ F_1 &= L_1 = \begin{bmatrix} I & O & O \\ O & I & O \\ O & H & I \end{bmatrix}, \qquad F_2 = L_2 = \begin{bmatrix} I & O & O \\ \Gamma & I & O \\ O & O & I \end{bmatrix}, \\ P &= \begin{bmatrix} \Sigma_{xx} & & \text{sym.} \\ O & \Sigma_{ee} & \\ O & O & \Sigma_{ff} \end{bmatrix}. \end{aligned} \tag{71}$$

This type of model now reduces the complex inverse of $(\mathbf{I} - \mathbf{A})$ into a much simpler series of sparse matrix multiplications. The other models with nested regressions (hierarchical factors in Equations 42–44 and path constraints in Equations 49 and 50) may be accomplished in similar fashion.

As a final example, let us consider that the general econometric nonrecursive model of RAM Equations 33 and 34 may be translated into an equivalent COSAN model given as

$$\begin{aligned} m &= 2, \qquad g_0' = [x'y'], \qquad g_1' = [x'y'], \qquad g_2' = [x'z'] \\ F_1 &= B_1^{-1} = \begin{bmatrix} I & O \\ O & (I - B) \end{bmatrix}, \qquad F_2 = L_2 = \begin{bmatrix} I & O \\ G & I \end{bmatrix}, \\ P &= \begin{bmatrix} \Sigma_{zz} & \text{sym.} \\ O & \Sigma_{xx} \end{bmatrix}. \end{aligned} \tag{72}$$

This COSAN model still requires an inversion at the first order, but this inverse problem is now well defined in a block-diagonal form that sparse matrix routines can more easily recognize and evaluate.

The major benefits of an algorithm with the features just defined are relatively easy to describe. First, this algorithm capitalizes on the simplicity of model specification using RAM $\mathbf{A}$ and $\mathbf{\Omega}$ parameters. These may be input either in compact elemental notation or block matrix form. Any alteration in model input, such as with canonically nested alternatives, requires only minimal additions or subtractions that will be transparent to the user. Second, the algorithm automatically creates the optimally efficient COSAN model for any RAM model by avoiding unnecessary inversion. This internal use of COSAN capitalizes on the general mathematical and algorithmic features of COSAN and permits the general calculation of estimates for any well defined loss function. Finally, the algorithm permits a direct translation back to the simple RAM notation for ease and flexibility of output [see McArdle & Horn (1981) for more details on the available algorithm].

The use of RAM theory as a modeling foundation leads to several other results of theoretical interest. For instance, unique estimates for $\mathbf{A}$ and $\mathbf{\Omega}$ parameters can only be obtained if the model is identified. General conditions for global model identification can now be simply described by writing a RAM model as

$$\mathbf{\Sigma} = \mathbf{F}\mathbf{E}^*\mathbf{\Omega}^*\mathbf{E}^{*\prime}\mathbf{F}', \tag{73}$$

where $\mathbf{E}^* = (\mathbf{I} - \mathbf{A})^{-1}\ \mathbf{T}$, $\mathbf{\Omega}^* = \mathbf{T}^{-1}\mathbf{\Omega}$, and $\mathbf{T}$ is any nonsingular transformation matrix (compare Bentler & Weeks, 1980, p. 295). This shows that any RAM model with an $\mathbf{F}$ prescribed by the variables in the model but with unknown $\mathbf{A}$ and $\mathbf{\Omega}$ parameters is globally identified if and only if $\mathbf{T} = \mathbf{I}$. This condition, of course, is quite difficult to verify for any problem of reasonable complexity. Therefore, although we have reduced the identification problem to a single $\mathbf{T}$-matrix form, the nonlinearity of $\mathbf{E} = (\mathbf{I} - \mathbf{A})^{-1}$ still requires detailed study of the alternative $\mathbf{T}$-matrix properties. Alternatively, it is still quite useful to note that any model is at least locally identified if $\mathbf{M} = \mathbf{H}$ (Equation 68) of full rank in large samples (McDonald & Krane, 1979).

These technical results have been written to be general for any given loss function f, but the statistical properties of the estimators for $\mathbf{A}$ and $\mathbf{\Omega}$ must be considered under the specific objective loss function defined by f, $\mathbf{G}$, and $\mathbf{M} = \mathbf{H}$ (Krane & McDonald, 1978; McDonald, 1978). In any case, the statistical problem considered here concerns making inferences about overall model fit from sample to population and determining the saliency of estimators. For any specific RAM model, we can write degrees of freedom (df) equal to

$$df = r(\mathbf{S}) - r(\mathbf{A}, \mathbf{\Omega}), \tag{74}$$

where $r(\mathbf{S}) = j(j+1)/2$, the rank of sample moment matrix $\mathbf{S}$, and $r(\mathbf{A}, \boldsymbol{\Omega})$ equals the rank of the matrix of first derivatives for $\mathbf{A}$ and $\boldsymbol{\Omega}$ with respect to $\boldsymbol{\Sigma}$ evaluated at some solution point. Under certain loss functions, such as the generalized least squares or maximum likelihood criteria (Browne, 1977; Lee & Jennrich, 1979), *asymptotic* properties of the estimators prove useful; the estimators are consistent, normally distributed, efficient, and scale invariant, and Nf_m is distributed as χ^2. This result yields the general comparison of other canonically nested models as a simple difference of

$$\chi_{\mathrm{d}}^2 = \chi_{\mathrm{h}}^2 - \chi_{\mathrm{a}}^2, \qquad \text{with} \quad df_{\mathrm{d}} = df_{\mathrm{h}} - df_{\mathrm{a}}, \tag{75}$$

where subscript h denotes a hypothesized model, subscript a denotes a more global alternative, and subscript d denotes the difference of these two models (Pesaran & Deaton, 1978). The comparison of any model against the trivial alternative where $\boldsymbol{\Sigma}$ is precisely equal to $\mathbf{S}$ with $df = j(j+1)/2$ yields a general statistical test for the global adequacy of model fit. The saliency of any individual parameter may now be determined from the calculation of standard errors for each parameter from $diag\ [2/N(\mathbf{H}^{-1})]$.

Many other useful indices that relate directly to goodness of fit and the general process of statistical model building (Saris et al., 1978; Sörbom, 1975) may be developed from RAM theory. For example, consider that the generalized least squares criterion for any RAM model may be explicitly written

$$f = 1/2\, tr[\{\mathbf{S}^{-1}(\mathbf{S} - \mathbf{F}(\mathbf{I} - \mathbf{A})^{-1}\boldsymbol{\Omega}(\mathbf{I} - \mathbf{A})^{-1\prime}\mathbf{F}')\}^2]. \tag{76}$$

We can now see that the function f may itself be elementally decomposed into residual elements, each of which is in turn based on sums of pathway products of $\mathbf{A}$ and $\boldsymbol{\Omega}$ elements that make up the $\mathbf{A}$ elements. In this way the weighted effect on any function value may now be calculated for each *individual* structural parameter. Another way to use this logic is to write an index of the model-fitting residuals,

$$g = (|\mathbf{F}(\mathbf{I} - \mathbf{A})^{-1}\boldsymbol{\Omega}(\mathbf{I} - \mathbf{A})^{-1\prime}\mathbf{F}'| - |\mathbf{S}|)/|\mathbf{S}|. \tag{77}$$

This index, based on the concept of a determinant as a measure of generalized variance, can be interpreted as an overall ratio of variance in the data explained by the model (Specht, 1975), and g can also be decomposed with elemental contributions of $\mathbf{A}$ and $\boldsymbol{\Omega}$ parameters. Similar concepts have been studied by Montanelli (1974) and Horn and Engstrom (1979), and associated statistical tests for these residual indicators have been studied by McDonald (1975). Of importance here may be the evaluation of the separate contributions of $\mathbf{A}$ and $\boldsymbol{\Omega}$ parameters, say in $|\boldsymbol{\Sigma}|$, and their relation to overall model fit. However, these issues are far too complex to be discussed in detail here (Horn & McArdle, 1980).

CONCEPTUAL REPRESENTATION

The mathematical development of RAM is clearly based on the previous treatments of Jöreskog (1973b), Wiley (1973), McDonald (1978), and Sörbom (1978). An especially heavy emphasis has been placed on some of the statistical and algorithmic possibilities (e.g., fixed vs. free parameters). In this sense, we recognize that RAM is not at all new, or even novel. Perhaps RAM is best considered as a consolidation of the basic ideas of the LISREL model, or the complete mathematical simplification of the COSAN model. However, we have also shown that this consolidation is illusory in many theoretical and practical ways. For example, it is clear that the current popularity of the LISREL model is in some part due to its theoretical development as a psychometrically based "measurement" model combined with an econometrically based "structural" model (Jöreskog & Sörbom, 1980). There is no doubt that the distinction between the measurement and structural components of a model is both practically and conceptually useful, especially when statistical issues of parameter estimation are considered. However, the algebraic basis of RAM clearly shows that these kinds of distinctions are not mathematically essential for full model development in terms of specification, estimation, evaluation, or even simulation (McArdle, 1980). Furthermore, these distinctions can breakdown when, for example, dealing with mean structures and/or complex model constraints (see Fig. 9.6).

The graphic development of RAM is clearly based on a combination of concepts in diagraph theory (Christofide, 1975; Harary, 1973) with those of traditional path-analytic diagrammatic representation (Wright, 1934, 1964; see also Li, 1975). Once again, most of the RAM ideas are not at all new, but represent a consolidation of traditional ideas. However, RAM graphics attempt to extend traditional path diagrams in two important ways. First, the RAM formulation requires the explicit definition of all undirected turning paths, even those associated with a single point (e.g., variances), whereas more traditional representations only require undirected turns between two sets of points (e.g., covariance or correlation). This explicit representation of all undirected slings is required simply because these slings are critical elements of the Σ associations. The second major difference between RAM graphics and traditional path diagrams is in the implicit representation of the residual variables $\mathbf{u}$. In RAM, latent residual variables with unit weights are not explicitly included, whereas in more traditional representations these are included, albeit not rigorously or consistently. Although it is possible to explicitly include these variables, as latent variables with fixed, unity paths (see Fig. 9.6c), this is unnecessary because it leads to redundancy in the associated algebra (see Equations 43 and 44). In general, the early work in path analysis did not entertain some possibilities that we wish to study with RAM, including models for covariances and means, as well as models with more complex latent common factor networks. Thus, the graphic alterations that we propose here are not arbitrary but are specifically designed to deal with these more general linear algebraic structures. Of course,

these new alterations also directly lead to the formal proof of the algebraic–graphic isomorphism.

The theoretical results on the RAM–COSAN–LISREL interchange are surprisingly complex; RAM, COSAN, and LISREL each may be considered as subsets of one another. An important feature of this interchange arises when we consider that the general RAM, COSAN, and LISREL models may each be written as more deeply subpartitioned sets of one another or themselves in an infinite series of translation (Hofstadter, 1979). This paradoxical interchange among general modeling foundations, first described by McArdle (1978, 1979b), and remarked on by Bentler and Weeks (1979), has important implications for structural modeling. For example, this demonstration on the equivalence of these models directly rejects earlier claims about one model being "more general" than another (Bentler, 1976; Bentler & Woodward, 1978; Weeks, 1978; cf. McDonald, 1978, 1979). Also, as can be seen from the RAM–COSAN interchange algorithm, each general modeling system, including LISREL, has potentially different and important theoretical and practical virtues. This emphasis on the important individual features of any modeling system is muddled, if not lost completely, in the search for a "most general" superstructure. Our emphasis here has been to show that RAM representation in Equations 1–15 is the only model that is based solely on the *necessary and sufficient* features required for a general linear structural equation system of the type we have described.

As we have attempted to illustrate, RAM theory also has much practical utility. For instance, in any specific application, RAM can simplify model specification, in terms of algebra or graphics, as well as provide a simplified statistical and algorithmic estimation and testing procedure. These, in turn, lead to several new results that have heretofore not been utilized. Alternatively, we have not pursued other practical algorithmic features for specific models (Wermuth, 1980; Wold, 1980), especially those related to computational graphics (Chachra, Prabhakar, & Moore, 1979; Heap, 1972), or the many other mathematical developments that require detailed study.

The conceptual basis of RAM, though decidedly less formal, is as critical as any other representation herein. Most certainly, the final form of Equation 15 or the isomorphic representation provided by Fig. 9.4 presents possibilities for many alternative abstract formulations. Among these, the one that seems most parallel to RAM theory is the "input/output" logic offered by a general systems philosophy (Rapoport, 1972). Under this logic, we define a "system" simply as "a set of variables and the relationships among them." We ignore several critical features of any specific system until these features prove necessary or useful in the conceptualization of that system. For instance, in this development we initially described only general features of variables until we wished to move toward a model for manifest moment structures, at which point it was necessary to distinguish between manifest and latent variables. Critically, this distinction among variables does not immediately demand any further restrictions on the kind of relationships possible,

such as only latent variables can be directed at manifest variables. It is conceivable that another practically useful distinction would be between endogenous (e.g., dependent) and exogenous (e.g., independent) system variables (compare Bentler & Weeks, 1980). However, once again, this would strictly require reference to a specific set of directed relationships that, from our general systems point of view, is unnecessarily artificial at this point in model development (Rozeboom, 1978). The logical extension of this system development is the expression of RAM in COSAN terms when estimation is at issue.

In many senses, RAM can also be considered as a general foundation of system *limits* on model conceptualization. As such, RAM theory provides a generally abstract *meta*-model from which more practical and theoretically interesting forms can emerge. In this presentation, we have chosen not to discuss philosophical notions of *causality*, although clearly these require further discussion (McArdle, 1980; Roozeboom, 1978; Wold, 1954) and have a clear relevance to our general systems perspective. Instead, as Fig. 9.4 shows, the general nature of RAM is just that of a relatively free-form *network* of interrelationships among points in space. The inclusion of the directed/undirected relationship distinction, along with the additional restrictions on pathway information flow throughout this network, denotes the importance of *action* or momentum between points in space. The inclusion of the manifest/latent variable distinction emphasizes the importance of the manifest moment structure as an empirical representation of the overall system action. In the recent psychometric literature, only Cattell (1978) has previously proposed such a general "reticule of interlocking influences" (p. 200; compare Fig. 9.4). To emphasize the importance of this theoretical conception, we have chosen the term *Reticular Action Meta-model*, with the mnemonic *RAM*, to represent both our algebra and our graphics.

A CONTEMPORARY VIEW OF THE RAM RULES

The previous six sections described the origins of the concepts that have come to be known as the RAM rules. At the time these ideas were first being circulated (1978–1981) I was excited by the possibility that I had found something simple about the complex world of general SEM. I was also unsure whether these rules would ever be useful and unsure about how they would be received. In looking back on this treatment, it is now clear that few changes have been required in the original RAM rules. From my point of view, the RAM rules have created both benefits and problems. I now conclude by taking a brief look at what has happened to the RAM rules during the last 25 years after I first wrote the previous sections.

The use of path diagrams has become a common device for the communication of both simple and complex structural models. There are so many applications where SEMs have been depicted as path diagrams that we can select many as

exemplars where they had not been used before (e.g., McDonald, 1985). On a more theoretical level, path-diagram graphics assisted in the development of several interesting SEM issues, including exploratory model search strategies (e.g., Glymour et al., 1987; Spirtes et al., 1993), the improved understanding of equivalent structural models (e.g., Lee & Hershberger, 1990; Luijben, 1991; MacCallum et al., 1993; Raykov & Marcoulides, 2002), and the clearer delineation of unambiguous causal inference (McDonald, 1997; Pearl, 2000).

In this same sense RAM suggested the requirement *to diagram every parameter in the model*. On the surface, this does not seem to be an unusual rule, and many researchers apparently believe they are following this rule. However, as witnessed by the current SEM literature, this part of the RAM rules has not yet caught on; most articles on path analysis still do not include all model parameters in their diagrams (McDonald & Ho, 2002). The publication standard of path models seems to have emerged as a mixture of diagramming techniques based on graphics used by Sewell Wright (1921, 1934, 1964, 1983) mixed with the treatment of unobservables in MIMIC models by Hauser and Goldberger (1971). To be sure, Sewell Wright anticipated most every aspect of current diagrams (McArdle & Aber, 1990), but he only made path diagrams as models of correlations. (As an aside, Hauser and Goldberger used circles for manifest variables and squares for latent variables.) However, as the models began to be based on covariances (due to equality constraints) and subsequently included means and intercepts, Wright's classic diagramming conventions prove inadequate in a fundamental way—incomplete diagrams do not allow the reader to count the parameters or to reproduce the model expectations using the tracing rules (Loehlin, 1998).

One oversight in most current published path diagrams is the lack of any self-slings (ω_{ii}) even though these are essential to the scaling of any predictor or latent variables and the subsequent interpretation of the other parameters. Many researchers do not label the residual disturbances or unique factor scores, and this makes it difficult to understand whether variances or deviation terms were fitted (and this can make a difference, see Equation 47). Only a few recent researchers include the means and intercepts in a fashion consistent with the tracing rules of cross-products (e.g., Equations 35–38), and some of these researchers do not include the average cross-product sling (labeled 1) on the constant (possibly due to the common misinterpretation of "slings as variances"). Some recent diagrams intended to reflect mixed-effect or multilevel models have drawn what is described as a "path into a path" (e.g., Muthén & Muthén, 2002, p. 245). Whereas this diagram may be useful in conveying the overall intention of the analysis, it is not a valid path diagram, in that it violates the rules for reproducing the expectations via tracing rules (for multilevel path diagrams, see McArdle & Hamagami, 1996).

The RAM rules are not fully documented in the major programs of SEM, such as LISREL or COSAN. This is no surprise, because these programs were developed using a different algebraic basis. However, the value of using RAM in LISREL programming had been demonstrated and used by McArdle (1978, 1979b, 1979c,

1980; Horn & McArdle, 1980) and subsequently observed by Graff and Schmidt (1982), among others. The use of RAM rules in computer programs without some provision for sparse matrix algebra slows down the calculation at each iteration (i.e., a patterned inverse is always used even if it is not needed; e.g., Equation 70). Nevertheless, whereas a RAM-based LISREL analysis would not be the fastest way for LISREL to do computer calculations, it turns out to be far easier to set up and debug for almost any problem (e.g., consider Fig. 9.1 as LISREL matrices).

In most recent versions of LISREL, a fixed "IZ" and a "No-X" model have been added as general options, and this makes RAM-type analysis far easier to use with the LISREL program (Jöreskog & Sörbom, 1999; McArdle, 2001). In one sense, the RAM rules do not exactly highlight key benefits of the standard LISREL approach, that is, "You may completely ignore five of the eight LISREL matrices here and still generate the same answer!" (e.g., Equations 55–56). Nevertheless, the more user-friendly version of the LISREL manual now recognize that a "No-X, No-ε" 3-matrix LISREL model can precisely match, iteration by iteration, any analysis that a full (now 12-matrix) LISREL model can accomplish. In the many complex applications where LISREL is most useful, the measurement–structural model separation is not the most critical feature, and the standard LISREL matrices have no substantive meaning (e.g., again see Fig. 9.1).

One clear benefit of the RAM rules is that they are not limited to any specific SEM program. As long as a SEM program permits the flexibility of a fixed $\mathbf{F}$, a patterned inverse of regression parameters $\mathbf{A}$, and a patterned symmetric $\mathbf{\Omega}$, this program can fit most any model (e.g., Fig. 9.1). This gives the user the flexibility to choose a program that is most appropriate for the specific problem at hand. In publications, it is possible to simply list all routines that produce similar results (and most do). These conceptual features continue to make RAM-based SEM programming an effective way to begin to use the fundamental and extensive features of excellent programs like LISREL and COSAN. My colleagues and I created our own computer program for model fitting based on the RAM–COSAN interface described here (e.g., RAMIT). The program we published was a subsection based on the tracing rules of path diagrams, RAMpath, by McArdle and Boker (1990; Boker et al., 2002).

Some problems are now outdated because several elegant new computer programs do emphazise RAM rules in their manuals. This started with the RAMONA program by Browne and Mels (1999), which was entirely based on the RAM algebra of Equations 1–15, and the R-SEM package by Fox (2002). Other programs with general matrix algebra have provisions for the RAM rules, including SAS PROC CALIS with the use of the RAMLIST (Hatcher, 1994) and the current Mx program by Neale et al. (1999). Mx is a freeware program that also includes a graphical user interface where a latent variable path model can be drawn and automatically turned into RAM matrices for model fitting and then turned back into presentation-style diagrams (following an algorithm devised by Boker; see Boker et al., 2002). Apparently the AMOS program (Arbuckle & Wotke, 1999) uses the

RAM rules and techniques, but the internal calculations of the AMOS program are not as transparent as those of the Mx program, and the path diagram output does not completely follow the RAM rules presented here (e.g., no self-loops are output, although this would be an easy alteration).

Most benefits of RAM have accrued from novel but practical applications. Our first set of applications included using SEM to deal with fairly complex developmental systems analyses, including *nonlinear constraints on age, time, and cohort* (Horn & McArdle, 1980; McArdle, 1981). In this work, our nonstandard use of "nonvariable" nodes anticipated the subsequent popularity of "phantom" variables by Rindskopf (1984; personal communication, 1986). This represented the beginning of our SEM re-interpretation of the then-new statistical models for dealing with incomplete data simply by writing different filter matrices ($\mathbf{F}^{(g)}$) for groups defined by the patterns of available data (using MAR assumptions; McArdle, 1994; McArdle & Anderson, 1990; McArdle & Hamagami, 1992). Based on the RAM rules, these complex incomplete data concepts were relatively easy to program and apply, and far easier than a standard LISREL approach (e.g., Allison, 1987).

A second set of applications came in the form of *novel behavioral genetics twin analyses* using existing SEM software (e.g., LISREL and COSAN; McArdle, 1986; McArdle, Connell, & Goldsmith, 1980; McArdle & Goldsmith, 1984, 1990). It seemed natural to show how RAM-based twin and family models could be expanded to include common factors, means and intercepts, and various forms of longitudinal data. This use of LISREL-based SEM was not the standard at the time, but appears now to be common (e.g., Henderson, 1992; McArdle & Prescott, 1996).

A third set of applications emerged when we ran across the classical *latent growth models* defined in SEM terms by Meredith and Tisak (1990). Based on RAM rules, these models were easy to represent as nonstandard but complete path diagrams for means and covariances that then could be fitted easily with any SEM software (McArdle, 1988; McArdle & Aber, 1990; McArdle & Epstein, 1987). The mixture of incomplete longitudinal data using simple filter matrices again proved to be a natural feature of RAM analyses (e.g., McArdle & Bell, 2000; McArdle & Hamagami, 1992; McArdle & Woodcock, 1997).

A final set of applications comes from our analyses of longitudinal data using *dynamic time-series models* for individuals and groups (McArdle, 1982; Nesselroade & McArdle, 1986; Nesselroade et al., 2002; Wood, 1988). These concepts led to the more recent representation of longitudinal data using latent difference scores (Hamagami & McArdle, 2000; McArdle, 2000; McArdle & Hamagami, 2001; McArdle & Nesselroade, 1994) and differential equations models (Boker, 2001). This accumulation of prior RAM-based work represents the building blocks for our recent dynamic analyses of incomplete longitudinal multivariate twin data based on models like Fig. 9.1.

Many of us have been able to teach SEM to students using these rules and have found that RAM simplifies many tedious tasks. We recognize that the majority

of SEM researchers still rely on the measurement–structural LISREL logic in presentations defined as "fundamental for beginners" (e.g., Bollen, 1988; Byrne, 1998; Mueller, 1996; Raykov & Marcoulides, 2000; Schumacker & Lomax, 1996). Alternatively, using the RAM rules, we can start with a graphic logic, sometimes termed the "RAM Game"(C.E. McArdle, Personal Communication, April, 1987): "Draw circles or squares, one-headed arrows, or two-headed slings." "If you can draw the diagram, you can automatically write the matrices." "Any program can generate the expectations." This starting point allows us to move on to the real hard questions of SEM, such as "What is your model?" This didactic aspect of the RAM rules should not be understated because it is often the novice who uses and appreciates RAM rules and the RAM Game. This approach also led to the development of new teaching tools, such as the automatic generation of SEM expectations and tracing rules using the new MAPLE and Mathematica programs.

Other benefits of RAM rules have been created by other researchers. Of course, many people now use RAM rules in their own presentations (e.g., Browne & Nesselroade, chap. 13, this volume; Loehlin, 1998; McDonald, 1995), and many have commented on their own approach to the RAM rules (e.g., "RAM is not wrong"; M. Browne, personal communication, October 2002). Indeed, Rod McDonald has extended this theory in several directions, including elegant studies in path analysis (McDonald, 1985, 1986, 1997; McDonald & Ho, 2002; McDonald, Parker, & Ishizuka, 1993) the creation of a "bi-level RAM" (Goldstein & McDonald, 1988; McDonald, 1994; McDonald & Goldstein, 1989), and the automatic generation of starting values for any SEM (Fox, 2002; McDonald & Hartman, 1992).

In this same sense, I hope I have demonstrated that RAM is a derivative work that stems from the fundamental flexibility of the original LISREL and COSAN concepts. As I look back on this work, I am not surprised that researchers who were invested in finding a more general system considered that this could only be done by expanding on these matrix representations (e.g., Bentler, 1976) and by creating new and complex programs for higher order systems (e.g., Weeks, 1978). Nor should it be any surprise that coming at the same problem from a different viewpoint—path graphics for SEM programs—could create an alternative view of the flexibility of SEM. I am grateful to those who have considered and improved on these ideas, so, after 25 years, I can use current vernacular and simply restate the simple principle, "RAM Rules."

ACKNOWLEDGMENTS

This research was initially supported by NSF-RIAS grant SER77-06935 and by NIA grant R01-AG02695. Special thanks go to colleagues who have contributed in ways too numerous to detail here, especially John L. Horn, Steven M. Boker, Roderick P. McDonald, Colin Fraser, Michael Browne, William Meredith, John R. Nes-

selroade, Fumiaki Hamagami, Kevin Grimm, Keith Widaman, and Albert Maydeu-Olivares for their advice and support. Requests for reprints should be sent to the author at jjm@virginia.edu reprints can also be downloaded at http://kiptron.psyc.virginia.edu.

REFERENCES

Circa 1981

Anderson, T. W., & Rubin, H. (1956). Statistical inference in factor analysis. In J. Neyman (Ed.): *Proceedings of the Third Berkeley Symposium on Mathematical Statistics and Probability* (pp. 111–150). Berkeley: University of California Press.

Bentler, P. M. (1976). Multistructure statistical model applied to factor analysis. *Multivariate Behavioral Research, 11*, 3–25.

Bentler, P. M., & Woodward, J. A. (1978). A head start reevaluation: Positive effects are not yet demonstrable. *Evaluation Quarterly, 2*, 493–510.

Bentler, P. M., & Weeks, D. G. (1979). Interrelations among models for the analysis of moment structures. *Multivariate Behavioral Research, 14*, 169–186.

Bentler, P. M., & Weeks, D. G. (1980). Linear structural equations with latent variables. *Psychometrika, 45*(3), 289–308.

Bock, R. D., & Bargmann. R. E. (1966). Analysis of covariance structures. *Psychometrika, 1966, 31*, 507–533.

Bock, R. D. (1975). *Multivariate statistical methods in behavioral research*. New York: McGraw-Hill.

Browne, M. W. (1977). Generalized least-squares estimators in the analysis of covariance structures. In D. J. Aigner & A. S. Goldberger (Eds.), *Latent variables in socio-economic models* (pp. 205–226). Amsterdam: North-Holland.

Cattell, R. B. (1978). *The scientific use of factor analysis in behavioral and life sciences*. New York: Plenum Press.

Chachra, V., Prabhakar, M. G., & Moore, J. M. (1979). *Applications of graph theory algorithms*. New York: North-Holland.

Christofide, N. (1975). *Graph theory: An algorithmic approach*. New York: Academic Press.

Dennis J. E., & Moré, J. J. (1977). Quasi-Newton methods, motivation and theory. *SIAM Review, 19*(1), 46–89.

Duff, I. S., & Reid, J. K. (1979). Some design features of a sparse matrix code. *ACM Transactions on Mathematical Software, 5*(1), 18–35.

Duncan, O. D. (1975). *Introduction to structural equation models*. New York: Academic Press.

Gustavson. F. G. (1978). Two fast algorithms for sparse matrices: multiplication and permuted transposition. *ACM Transactions on Mathematical Software, 4*(3), 250–269.

Guttman, L. (1957). Simple proofs of relations between the communality problem and multiple correlation. *Psychometrika, 22*, 147–158.

Hanushek, E. A., & Jackson, J. E. (1977). *Statistical methods for social scientists*. New York: Academic Press.

Harary. F. (1973). *New directions in the theory of graphs*. New York: Academic Press.

Harary, F., Norman, R. Z., & Cartwright, D. (1965). *Structural models: An introduction to the theory of directed graphs*. New York: Wiley.

Heise, D. R. (1975). *Causal analysis*. New York: Wiley Press.

Heap, B. R. (1972). The production of graphs by computer. In R. C. Read (Ed.), *Graph Theory and Computing* (pp. 10–15). New York: Academic Press.

Hofstadter, D. R. (1979). *Goedel, Escher, Bach: An Eternal Golden Braid*. NY: Basic Books.

Horn, J. L., & Engstrom, R. (1979). Cattell's scree test in relation to Bartlett's chi-square test and other observations on the number of factors problem. *Multivariate Behavioral Research, 14*, 283–300.

Horn, J. L., & McArdle, J. J. (1980). Perspective on mathematical and statistical model building (MASMOB) in aging research. In L. W. Poon (Ed.), *Aging in the 1980's* (pp. 503–541). Washington. D.C.: American Psychological Association.

Jennrich, R. I. (1977). Stepwise Regression. In K. Enslein, I. Ralston & H. W. Wilf (Eds.), *Statistical methods for digital computers-Vol III* (pp. 18–25). New York: Wiley Press.

Jöreskog, K. G. (1970). A general method for the analysis of covariance structures. *Biometrika, 57*, 239–251.

Jöreskog. K. G. (1971). Simultaneous factor analysis in several populations. *Psychometrika, 36*(4). 409–426.

Jöreskog, K. G. (1973a). Analysis of covariance structures. In A. S. Goldberger & O. D. Duncan (Eds.), *Structural equation models in the social sciences* (pp. 53–68). New York: Seminar Press.

Jöreskog, K. G. (1973b). A general method for estimating a linear structural equation system. In A. S. Goldberger & O. D. Duncan (Eds.), *Structural equation models in the social sciences* (pp. 85–112). New York: Seminar Press.

Jöreskog, K. G. (1971). Simultaneous factor analysis in several populations. *Psychometrika, 36*, 409–426.

Jöreskog. K. G., & Sörbom, D. (1980). *Simultaneous analysis of longitudinal data from several cohorts. (Research Report 80-5)*. Uppsala, Sweden; University of Uppsala, Department of Statistics.

Jöreskog, K. G., & Sorbom, D. (1979). *Advances in factor analysis and structural equation models*. Cambridge, MA: Abt Assoc.

Jöreskog, K. G., & Sörbom, D. (1978). LISREL-IV: *Analysis of linear structural relationships by the method of maximum likelihood*. Chicago: National Educational Resources.

Kenny, D. A. (1979). *Correlation and causality*. New York: Wiley.

Krane, W. R., & McDonald, R. P. (1978). Scale invariance and the factor analysis of correlation matrices. *British Journal of Mathematical Statistical Psychology, 31*, 218–228.

Lee, S. Y., & Jennrich, R. I. (1979). A study of algorithms for covariance structure analysis with specific comparisons using factor analysis. *Psychometrika, 44*, 99–113.

Li, C. C. (1975). *Path analysis*. Pacific Grove, CA: Boxwood Press.

McArdle, J. J. (1978). *A structural view of structural models*. Paper presented at the Winter Workshop on Latent Structure Models Applied to Developmental Data. University of Denver.

McArdle, J. J. (1979). *A SYSTEMatic view of structural equation modeling*. Paper presented at the Psychometric Society Annual Meeting, Monterey, California.

McArdle, J. J. (1979). *Reticular Analysis Modeling (RAM) theory: the simplicity and generality of structural equations*. Paper presented at the American Psychological Association Annual Meeting, New York.

McArdle, J. J. (1979). The development of general multivariate software. *Proceedings of the Association for the Development of Computer-Based Instructional Systems* (pp. 824–862). Akron, Ohio: University of Akron Press.

McArdle, J. J. (1980). Causal modeling applied to psychonomic systems simulation. *Behavior Research Methods and Instrumentation, 12*, 193–209.

McArdle, J. J., & Horn, J. L. (1981). *MULTITAB 81: Computer programs for multivariate data analysis*. Department of Psychology, University of Denver.

McArdle, J. J., & McDonald, R. P. (1980). *A note on some properties of the Reticular Action Model for moment structures*. Department of Psychology, University of Denver.

McDonald, R. P. (1969). A generalized common factor analysis based on residual covariance matrices of prescribed structure. *British Journal of Mathematical Statistical Psychology, 22*, 149–163.

McDonald, R. P. (1975). A note on Rippe's test of significance in common factor analysis. *Psychometrika, 40*(1), 117–119.

McDonald, R. P. (1978) A simple comprehensive model for the analysis of covariance structures. *British Journal of Mathematical Statistical Psychology, 31*, 59–72.

McDonald, R. P. (1979). The structural analysis of multivariate data: A sketch of a general theory. *Multivariate Behavioral Research, 1979, 14*, 21–28.

McDonald, R. P. (1980). A simple comprehensive model for the analysis of covariance structures: some remarks on application. *British Journal of Mathematical Statistical Psychology, 33*, 161–183.

McDonald, R. P., & Krane, W. R. (1979). A Monte Carlo study of local identifiability and degrees of freedom in the asymptotic likelihood ratio test. *British Journal of Mathematical Statistical Psychology, 32*, 121–132.

McDonald, R. P., & Swaminathan, H. (1972). *Structural analyses of dispersion matrices based on a very general model with a rapidly convergent procedure for the evaluation of parameters*. Informal publication of M.E.C.A., O.I.S.E, Toronto.

Montanelli, R. G. (1974). The goodness of fit of the maximum-likelihood estimation procedure in factor analysis. *Educational and Psychological Measurement, 34*, 547–562.

Pesaran, M. R., & Deaton, A. S. (1978). Testing non-nested nonlinear regression models. *Econometrika, 46*, 677–694.

Rapoport, A. (1972). The uses of mathematical isomorphism in general systems theory. In G. J. Klir (Ed.), *Trends in general systems theory* (pp. 1–10). New York: Wiley.

Rozeboom, W. W. (1978). *General Linear Dynamic Analysis (GLDA)*. Paper presented at the Winter Workshop on Latent Structure Models Applied to Developmental Data. University of Denver.

Saris, W. E., dePijper, M., & Zewaart, P. (1978). Detection of specification errors in linear structural equation models. In K. F. Schuessler (Ed.), *Sociological methodology* (pp. 187–197). San Francisco: Jossey Bass.

Schmid, J., & Leiman, J. (1957). The development of hierarchical factor solutions. *Psychometrika, 22*, 53–61.

Sörbom. D. (1974). A general method for studying differences in factor means and factor structure between groups. *British Journal of Mathematical Statistical Psychology, 27*, 229–239.

Sörbom, D. (1978). An alternative to the methodology for analysis of covariance. *Psychometrika, 43*, 381–396.

Spearman, C. (1904). General intelligence objectively determined and measured. *American Journal of Psychology, 15*, 201–293.

Specht, D. A. (1975). On the evaluation of causal models. *Social Science Research, 4*, 113–133.

Thurstone, L. L. (1947). *Multiple factor analysis*. Chicago: University of Chicago Press.

van Driel, O. P. (1978). On various causes of improper solutions in maximum likelihood factor analysis. *Psychometrika, 43*(2), 225–243.

Weeks, D. G. (1978). *A second order structural equation model of ability*. Paper presented at the Winter Workshop on Latent Structure Models Applied to Developmental Data. University of Denver.

Wermuth, N. (1980). Linear recursive equations. covariance selection, and path analysis. *Journal of the American Statistical Association, 75*, 963–972.

Wiley, D. E. (1973). The identification problem for structural equation models with unmeasured variables. In A. S. Goldberger & O. D. Duncan (Eds), *Structural equation models in the social sciences* (pp. 69–83). New York: Seminar Press.

Wold, H. O. A. (1954). Causality and econometrics, *Econometrika, 22*, 162–177.

Wold, H. O. A. (1960). A generalization of causal chain models. *Econometrika, 28*, 443–463.

Wold, H. O. A. (1980). Soft modelling: Intermediate between traditional model building and data analysis. *Mathematical Statistics, 6*, 333–346.

Wright, S. (1921). Correlation and causation. *Journal of Agricultural Research, 20*, 557–585.

Wright, S. (1934). The method of path coefficients. *Annals of Mathematical Statistics, 5*, 161–215.

Wright, S. (1964). The interpretation of multivariate systems. In O. Kempthorne, T. A. Bancroft, J. W. Gowen & J. L. Lush (Eds.), *Statistics and mathematics in biology* (pp. 100–112). New York: Hafner.

Circa 2004

Allison, P. D. (1987). Estimation of linear models with incomplete data. In C.C. Clogg (Ed.). *Sociological Methodology 1987*, San Fran.: Jossey-Bass, 71–103.

Arbuckle, J. L., & Wotke, W. (1999). AMOS 4.0 User's Guide. Chicago: Smallwaters.

Boker, S. M. (2001). Differential models and differential structural equation modeling of intraindividual variability. In L. M. Collins & A. G. Sayer (Eds.), *New methods for the analysis of change* (pp. 5–27). Washington DC: APA.

Boker, S. M., McArdle, J. J., & Neale, M. (2002). An algorithm for the hierarchical organization of path diagrams and calculation of components of expected covariance. *Structural Equation Modeling, 9*, 174–194.

Bollen, K. A. (1989). *Structural equations with latent variables*. New York: Wiley.

Browne, M. W., & Mels, G. (1999). Path Analysis (RAMONA). In SYSTAT 10, Statistics II [Computer software and manual] (pp. II-233–II-291). Chicago: SPSS Inc.

Byrne, B. (1998). *Structural equation modeling with LISREL, PRELIS, and SIMPLIS. Basic concepts, applications, and programming*. Mahwah, NJ: Erlbaum.

Fox, J. (2002). *SEM (Structural Equation Modeling) for R. CRAN* (the Comprehensive R Archive Network). R Foundation for Statistical Computing, Vienna University of Technology (Technische Univasitat Wien).

Glymour, C., Scheines, R., Spirtes, P., & Kelly, K. (1987). Discovering causal structure: Artifical intelligence. philosophy of science, and statistical modeling. New York: Academic Press.

Goldstein, H., & McDonald, R. P. (1988). A general model for the analysis of multilevel data. *Psychometrika, 53*, 455–467.

Graff. J., & Schmidt, P. (1982). A general model for the decomposition of effects. In K. G. Jöreskog & H. Wold (eds.), *Systems under indirect observation: Causality, structure, prediction* (pp. 131–148). North-Holland.

Hamagami, F., & McArdle, J. J. (2000). Advanced studies of individual differences linear dynamic models for longitudinal data analysis. In G. Marcoulides & R. Schumacker (Eds.), *Advanced structural equation modeling: Issues and techniques* (pp. 203–246). Mahwah, NJ: Erlbaum.

Hatcher, L. (1994). *A step-by-step approach to using the SAS system for factor analysis and structural equation modeling*. Cary, NC: SAS Publishing.

Hauser, R. M., & Goldberger, A. S. (1971). The treatment of unobservable variables in path analysis. In H. L. Costner (Ed.), *Sociological Methodology, 1971* (pp. 81–117). San Francisco: Jossey-Bass.

Hayduk, L. (1987). *Structural equation modeling with LISREL: Essentials and Advances*. Baltimore: John Hopkins University Press.

Henderson, N. D. (1982). Human behavior genetics. *Annual Review of Psychology, 33*, 403–440.

Jöreskog, K. G., & Sörbom, D. (1999). LISREL 8.30: *LISREL 8: Structural equation modeling with the SIMPLIS command language*. Hillsdale, NJ: Scientific Software International.

Lee, S., & Hershberger, S. (1990). A simple rule for generating equivalent models in covariance structure modeling. *Multivariate Behavioral Research, 25*, 313–334.

Luijben, T. C. W. (1991). Equivalent models in covariance structure modeling. *Psychometrika, 56*, 653–665.

Loehlin, J. C. (1987, 1998). *Latent variables: an introduction to factor, path, and structural analysis*. Mahwah, NJ: Laurence Erlbaum Associates.

MacCallum, R. C., Wegener, D. T., Uchino, B. N., & Fabrigar, L. R. (1993). The problem of equivalent models in applications of covariance structure analysis. *Psychological Bulletin, 114*, 185–199.

McArdle, J. J. (1982). Structural Equation Modeling of an Individual System: Preliminary Results from "A Case Study of Episodic Alcoholism." Department of Psychology, University of Denver, Research report to the *National Institute of Alcoholism and Alcohol Abuse*.

McArdle, J. J. (1988). Dynamic but structural equation modeling of repeated measures data. In J. R. Nesselroade & R.B. Cattell (Eds.), *The Handbook of Multivariate Experimental Psychology, Volume 2* (pp. 561–614). New York, Plenum Press.

McArdle, J. J. (1986). Latent variable growth within behavior genetic models. *Behavior Genetics, 16*(1), 163–200.

McArdle, J. J. (1994). Structural factor analysis experiments with incomplete data. *Multivariate Behavioral Research, 29*(4), 409–454.

McArdle. J. J. (2001). A latent difference score approach to longitudinal dynamic structural analyses. In R. Cudeck, S. du Toit, & D. Sörbom (Eds.), *Structural Equation Modeling: Present and future* (pp. 342–380). Lincolnwood, IL: Scientific Software International.

McArdle. J. J.. & Aber, M. S. (1990). Patterns of change within latent variable structural equation modeling. In A. von Eye (Ed.), *New Statistical Methods in Developmental Research* (pp. 151–224). New York: Academic Press.

McArdle. J. J.. & Anderson, E. (1990). Latent variable growth models for research on aging. In J. E. Birren & K.W. Schaie (Eds.), *The Handbook of the Psychology of Aging* (pp. 21–43). New York: Plenum Press.

McArdle. J. J., & Bell, R. Q. (2000). Recent trends in modeling longitudinal data by latent growth curve methods. In Little, T. D., Schnabel, K. U., & Baumert, J. (Eds.), *Modeling longitudinal and multiple-group data: practical issues, applied approaches, and scientific examples* (pp. 69–108). Mahwah, NJ: Erlbaum.

McArdle, J. J., & Boker, S. M. (1990). *RAMpath: A Computer Program for Automatic Path Diagrams*. Hillsdale, NJ: Lawrence Erlbaum Publishers.

McArdle. J. J., Connell, J.. & Goldsmith, H. H. (1980). Latent variable approaches to measurement, structure, longitudinal stability, and genetic influences: Preliminary results from the study of behavioral style. *Behavior Genetics, 10*, 609.

McArdle. J. J., & Epstein. D. B. (1987). Latent growth curves within developmental structural equation models. *Child Development, 58*(1), 110–133.

McArdle, J. J., & Goldsmith, H. H. (1984). Multivariate biometric models of the WAIS. *Behavior Genetics, 14*, 609.

McArdle, J. J., & Goldsmith, H. H. (1990). Some alternative structural equation models for multivariate biometric analyses. *Behavior Genetics, 20*(5), 569–608.

McArdle, J. J., & Hamagami, E. (1992). Modeling incomplete longitudinal and cross–sectional data using latent growth structural models. *Experimental Aging Research, 18*(3), 145–166.

McArdle, J. J., & Hamagami, F. (1996). Multilevel models from a multiple group structural equation perspective. In G. Marcoulides & R. Schumacker (Eds.), *Advanced Structural Equation Modeling Techniques* (pp. 89–124). Hillsdale, NJ: Erlbaum.

McArdle, J. J., & Hamagami, F. (2001). Linear dynamic analyses of incomplete longitudinal data. In L. Collins & A. Sayer (Eds.), *Methods for the Analysis of Change* (pp. 137–176). Washington, DC: APA Press.

McArdle, J. J., & Hamagami, F. (2003). Structural equation models for evaluating dynamic concepts within longitudinal twin analyses. *Behavior Genetics, 33*(3), 137–159.

McArdle. J. J., & McDonald, R. P. (1984). Some algebraic properties of the Reticular Action Model for moment structures. *British Journal of Mathematical & Statistical Psychology, 37*(2), 234–251.

McArdle, J. J., & Nesselroade, J. R. (1994). Structuring data to study development and change. In S. H. Cohen & H. W. Reese (Eds.), *Life-Span Developmental Psychology: Methodological Innovations* (pp. 223–268). Hillsdale, NJ.: Erlbaum.

McArdle, J. J., & Prescott, C. A. (1996). Contemporary models for the biometric genetic analysis of intellectual abilities. In D. P. Flanagan, J. L., Genshaft. & P. L. Harrison (Eds.). *Beyond traditional intellectual assessment: Contemporary and emerging theories, tests and issues* (pp. 403–436). New York: Guilford Press.

McArdle. J. J., & Woodcock, J. R. (1997). Expanding test–rest designs to include developmental time–lag components. *Psychological Methods, 2*(4), 403–435.

McDonald, R. P. (1985). *Factor Analyses and Related Methods*. Mahwah, NJ: Lawrence Erlbaum Associates, Inc.

McDonald, R. P. (1986). Describing the elephant: Structure and function in multivariate data. *Psychometrika, 41*,(4) 513–534.

McDonald, R. P. (1994). The bilevel reticular action model for path analysis with latent variables. *Sociological Methods and Research, 22*, 399–413.

McDonald, R. P. (1997). Haldane's lungs: a case study in path analysis. *Multivariate Behavioral Research, 32*, 1–38.

McDonald, R. P., & Hartman, W. M. (1992). A procedure for obtaining initial values of parameters in the RAM model. *Multivariate Behavioral Research, 27*, 57–76.

McDonald, R. P., & Goldstein, H. (1989). Balanced versus unbalanced designs for linear structural relations in two-level data. *British Journal of Mathematical & Statistical Psychology, 42*, 215–232.

McDonald, R. P., & Ho. M. R. (2002). Principles and practice in reporting structural equation analyses. *Psychological Methods, 7*(1), 64–82.

McDonald, R. P., Parker, P. M., & Ishizuka, T. (1993). A scale-invariant treatment for recursive path models. *Psychometrika, 58*, 431–443.

Meredith, W., & Tisak, J. (1990). Latent curve analysis. *Psychometrika, 55*, 107–122.

Mueller, R. O. (1996). *Basic principles of structural equation modeling: An introduction to LISREL and EQS*. New York: Springer-Verlag.

Muthén, L. K., & Muthén, B. O. (1998). *Mplus, the comprehensive modeling program for applied researchers user's guide*. Los Angeles, CA: Muthén & Muthén.

Neale, M. C., Boker, S. M.. Xie. G., & Maes, H. H. (1999). Mx: Statistical Modeling. Box 126 MCV. Richmond, VA: Department of Psychiatry.

Nesselroade, J. R., & McArdle, J. J. (1986). Multivariable causal modeling in alcohol use research. *Social Biology, 33*(4), 272–296.

Nesselroade, J. J., McArdle, J. J., Aggen, S. H., & Meyers, J. (2001). Dynamic factor analysis models for multivariate time series analysis. In D. M. Moskowitz & S. L. Hershberger (Eds.), *Modeling individual variability with repeated measures data: Advances & Techniques* (pp. 233–266). Mahwah. New Jersey: Erlbaum.

Pearl, J. (2000). *Causality: Models Reasoning, and Inference*. Cambridge: Cambridge University Press.

Raykov, T., & Marcoulides, G. A. (2000). *A first course in structural equation modeling*. Mahwah, NJ: Erlbaum.

Raykov, T., & Marcoulides, G. A. (2002). Can there be infinitely many models equivalent to a given structural equation model? *Structural Equation Modeling, 8*, 142–149.

Rindskopf, D. (1984). Using phantom and imaginary latent variables to parameterize constraints in linear models. *Psychometrika, 49*, 37–47.

Schumacker, R. E., & Lomax, R. G. (1996). *A beginner's guide to structural equation modeling*. Mahwah, NJ: Erlbaum.

Spirites, P., Glymour, C., & Scheines, R. (1993). *Causation, Prediction, and Search*. New York: Springer-Verlag.

Wood. P. K. (1988). Dynamic factor analysis as reformulated under the Reticular Action Model. Paper presented at the annual meetings of the *American Educational Research Association*, New Orleans, LA.

Wright, S. (1983). Path analysis in epidemiology: A critique. *Amer. Jour. Human Genetics, 35*, 757–768.