DOCUMENT RESUME

ED 283 850 TM 870 369

AUTHOR Wothke, Warner
TITLE Multivariate Linear Models of the Multitrait-Multimethod Matrix.
PUB DATE 25 May 87
NOTE 48p.; Paper presented at the Annual Meeting of the American Educational Research Association (Washington, DC, April 20-24, 1987).
PUB TYPE Speeches/Conference Papers (150) -- Reports - Research/Technical (143)

EDRS PRICE MF01/PC02 Plus Postage.
DESCRIPTORS *Analysis of Covariance; Correlation; Estimation (Mathematics); *Factor Analysis; Factor Structure; Goodness of Fit; *Mathematical Models; Measurement Techniques; *Multitrait Multimethod Techniques; Statistical Studies
IDENTIFIERS *Confirmatory Factor Analysis; Congeneric Tests; Covariance Structure Models; LISREL Computer Program; Trait Factor Approach

ABSTRACT

Several multivariate statistical methodologies have been proposed to ensure objective and quantitative evaluation of the multitrait-multimethod matrix. The paper examines the performance of confirmatory factor analysis and covariance component models. It is shown, both empirically and formally, that confirmatory factor analysis is not a reliable method for simultaneous estimation of trait and method factors. The poor performance is due to an inherent rotational indeterminacy common to all factor analytic models of trait and method effects. Covariance component analysis, on the other hand, shows a more parsimonious parameterization of general, trait, and method variation in the multitrait-multimethod matrix and is therefore typically unaffected by rotational indeterminacies. The performance with 23 empirical multitrait-multimethod correlation matrices was also found satisfactory. (Author)

Multivariate linear models of the multitrait-multimethod matrix

Werner Wothke*
Northwestern University
Department of Psychology
Evanston, IL 60201

May 25, 1987

Abstract

Several multivariate statistical methodologies have been proposed to ensure objective and quantitative evaluation of the multitrait-multimethod matrix. The paper examines the performance of confirmatory factor analysis and covariance component models. It is shown, both empirically and formally, that confirmatory factor analysis is *not* a reliable method for simultaneous estimation of trait and method factors. The poor performance is due to an inherent rotational indeterminacy common to all factor analytic models of trait *and* method effects. Covariance component analysis, on the other hand, shows a more parsimonious parameterization of general, trait, and method variation in the multitrait-multimethod matrix and is therefore typically unaffected by rotational indeterminacies. The performance with 23 empirical multitrait-multimethod correlation matrices was also found satisfactory.

1 Qualitative foundations of the multitrait-multimethod approach

Pivotal to the arguments in the paper is the notion that method effects in behavioral research are (a) sizable, (b) undesirable, (c) products of many "potential influences at several levels of abstraction" (Fiske, 1982, p. 82), and that (d) "we have only *other invalid measures* against which to validate our tests; we have no 'criterion' to check them against" (Campbell, 1969, p. 15). The size of methods effects in individual measurements *cannot* be exactly determined in a platonic sense, but method dependence can be assessed in a crude sense

*R.D. Bock, D.W. Fiske, K.G. Jöreskog, and D. Rindskopf contributed valuable suggestions and questions.

when the measurements change with the assessment methods. The question is how trait validity may be assessed without having to know the exact nature of method disturbances beforehand.

This paper conceives of *traits* as constructs relating an unobservable magnitude to differences among observable units of measurement. For simplification, it is assumed that traits and their indicators are linearly related.

In their well-known paper, Campbell & Fiske (1959) proposed the multitrait-multimethod (MTMM) matrix format as a device to study trait validity across different assessment methods. The MTMM matrix shows a crossed measurement design based on a simple rationale: Traits (i.e., latent quantitative characteristics of the research units) are universal, equally manifest over a variety of situations and detectable with a variety of methods. Most importantly, traits should not change just because different assessment methods are used. Hence, if there are m multiple sets of measures of t traits, each utilizing a different method of assessment, and if the methods indeed produce equivalent measurements, then the resulting covariance matrix takes the form

$$\begin{aligned}\Sigma_{(mt\times mt)} &= \mathbf{1}_{(m\times m)} \otimes \Sigma_{\tau(t\times t)} + \mathrm{diag}(\theta_{(1,1)}, \theta_{(1,2)}, \ldots, \theta_{(m,t)})_{(mt\times mt)} \\ &= \left(\begin{array}{c|c|c} \Sigma_\tau & \ldots & \Sigma_\tau \\ \hline \vdots & \ddots & \vdots \\ \hline \Sigma_\tau & \ldots & \Sigma_\tau \end{array}\right) + \mathrm{diag}(\theta_{(1,1)}, \theta_{(1,2)}, \ldots, \theta_{(m,t)})_{(mt\times mt)} \quad (1)\end{aligned}$$

where

Σ is the $mt \times mt$ covariance matrix among all measures,

$\mathbf{1}$ is a $m \times m$ matrix of unit entries in all its elements,

$\otimes$ symbolizes the Kronecker product operator (cf. Bock, 1975),

Σ_τ is the covariance matrix among trait measures within each method, and

$\mathrm{diag}(\theta_{(1,1)}, \theta_{(1,2)}, \ldots, \theta_{(m,t)})$ is the diagonal matrix of uncorrelated uniqueness components of the mt measures.

The model described by Equation (1) is reasonable only if all measurements are made on the same scale. This is equivalent to the psychometric concept of τ-equivalent measurement (Lord & Novick, 1968). In the behavioral and social sciences, where very often diverse methods like test scores, behavioral observations, and one-item ratings are compared, such strong scale assumptions are usually not warranted. Scale information is typically regarded as arbitrary or of little interest and, since Spearman's days, the social sciences have had a tradition of analyzing *correlation* matrices, effectively neglecting information due to the original scale of measurement. For these reasons, some terminology designed to describe valid measurement in more general terms based on correlation patterns is preferable to the strict formulation of Equation (1).

Campbell & Fiske proposed several qualitative criteria to judge convergent and discriminant validity. These criteria are quite popular and appear to be rigorous, but can be shown to be not quite adequate in borderline cases (Wothke, 1984) and, because of their complete lack of any statistical basis, need to be replaced by quantitative rules (see, e.g.: Althauser, 1974; Althauser & Heberlein, 1970; Althauser, Heberlein & Scott, 1971).

Confirmatory factor analysis (Jöreskog, 1966, 1977) and *covariance component analysis* (Bock, 1960; Bock & Bargmann, 1966; Wiley, Schmidt & Bramble, 1973) are two quantitative approaches with potential application to multitrait-multimethod analysis. Both models are realizations of the multivariate linear model and are embedded in an abundance of statistical theory. Otherwise, they are structurally distinct and derive from different statistical traditions: Confirmatory factor analysis is rooted in the psychometric tradition of validity theory, as outlined by Lord & Novick (1968); covariance component analysis is a multivariate generalization of random effects analysis of variance, based on R.A. Fisher's work.

2 Confirmatory Factor Analysis of the Multitrait-Multimethod Matrix

Confirmatory factor analysis (CFA) of the MTMM matrix was first proposed by Jöreskog (1966, 1977). Employing essentially the same maximum-likelihood estimation techniques as Lawley's (1940) exploratory factor analysis, confirmatory factor analysis is commonly characterized by additional equality restrictions imposed on estimated factor loadings, factor variances and covariances, and on unique components of the measured variables. A computer program for CFA is available in LISREL-6 (Jöreskog & Sörbom, 1986). Assuming multivariate normality of factor space and measurement errors, maximum-likelihood χ^2 tests among nested models can be performed. In cases of non-normality, the more recent work by Browne (1984b) seems promising, using a weighted least-squares estimation approach.

Factor analysis decomposes the $n \times p$ data matrix $\mathbf{X}$ of p of measures on n units into an $n \times k$ matrix Ξ of a lesser number k of latent factors:

$$\mathbf{X} = \Xi\Lambda' + \mathbf{E}, \tag{2}$$

where Λ is the $p \times k$ matrix of partial regression coefficients of observed measures regressed onto the latent factors and $\mathbf{E}$ is the matrix of unique, uncorrelated components. Expressing the population covariance matrices of $\mathbf{X}$ as Σ_x, of Ξ as Φ, and of $\mathbf{E}$ as Θ, respectively, Equation 2 implies the covariance representation

$$\Sigma_x = \Lambda\Phi\Lambda + \Theta. \tag{3}$$

Over the years, different types of confirmatory factor models have been proposed for the multitrait-multimethod matrix. They can be characterized by how many latent factors are modeled; whether these factors are thought to describe trait variance, method variance, or both; and whether the correlation structure among these methods is free or restricted. Occasionally, models with correlated uniqueness coefficients were also applied (e.g., Stacy et al., 1985), but the present paper retains the classic factor analytic notion that unique components are uncorrelated. With this one restriction, all CFA models of the MTMM matrix may be described by particular restriction patterns imposed on the Λ and Φ matrices.

2.1 Trait-only factor analysis

The simplest factor analytic models require that all common variation among measures is due to the latent trait factors and that no covariation is due to the assessment methods. Different traits may be correlated. Such a trait-only model, which specifies that each measure assesses exactly one trait factor (Jöreskog, 1971, 1978; Schmitt, 1978; Werts, Jöreskog, & Linn, 1972; Werts & Linn, 1970) shows the properties of *congeneric* measurement. When measures are ordered by traits within measures, the factor loading matrix for a 3 × 3 MTMM design takes the form

$$\Lambda_\tau = \begin{pmatrix} \lambda_{1,\tau_1} & 0 & 0 \\ 0 & \lambda_{2,\tau_2} & 0 \\ 0 & 0 & \lambda_{3,\tau_3} \\ \lambda_{4,\tau_1} & 0 & 0 \\ 0 & \lambda_{5,\tau_2} & 0 \\ 0 & 0 & \lambda_{6,\tau_3} \\ \lambda_{7,\tau_1} & 0 & 0 \\ 0 & \lambda_{8,\tau_2} & 0 \\ 0 & 0 & \lambda_{9,\tau_3} \end{pmatrix} \tag{4}$$

and the matrix of factor intercorrelations is obtained by

$$\Phi_\tau = \begin{pmatrix} 1.0 & \text{(symm.)} & \\ \phi_{\tau_2,\tau_1} & 1.0 & \\ \phi_{\tau_3,\tau_1} & \phi_{\tau_3,\tau_2} & 1.0 \end{pmatrix}. \tag{5}$$

The factor structure is non-overlapping and oblique. All zero entries in Λ_τ and all diagonal entries in Φ_τ are fixed (predetermined) parameters, the nine symbolic parameters λ_{i,τ_j} and the three symbolic parameters $\Phi_{\tau_j,\tau_{j'}}$ are estimated from the data. Also estimated are the nine uniqueness coefficients in the diagonal of Θ.

The Campbell-Fiske criteria, applied to the trait-only model, appear as inequality constraints on factor loadings and factor intercorrelations. Two of the criteria, are rather useful, yet, because Campbell & Fiske tacitly assumed homogeneous reliabilities, the other two criteria imply a somewhat unintelligible trade-off between the boundary conditions for factor loadings and factor correlations (Wothke, 1984).

The most interesting question for the applied researcher is how well the congeneric trait-only model can be used to describe and analyze empirical MTMM matrices. Critical indicators for model performance are

Identification—uniqueness of the model parameter estimates. The parameters are not identified when *different sets of values* for one or more estimates result in the *same model covariance matrix* $\hat{\Sigma}$. A trivial case of non-identification occurs when the model has more parameters than elements in the model covariance matrix. Further issues concerning the identification problem with MTMM factor analysis are discussed in Wothke (1984).

Convergence—numerical evaluation of the parameter estimates. Maximum-likelihood parameter estimates for the covariance structure models considered here do not have a closed-form solution. Estimates must be obtained iteratively, using, for example, the Fletcher-Powell or Newton-Raphson algorithms implemented in the LISREL program. Both numerical methods are generally efficient, but can fail to converge to a final solution, either because the number of iterations will exceed the present program limit of 250 or because the process will actually diverge. Common reasons for non-convergence are (a) starting values chosen too far from the final solution, (b) flat maxima or ridges in the likelihood surface, and (c) singularity of the information matrix in the vicinity of the solution. The latter two cases indicate poor model properties.

Admissibility—the Fisherian estimation methods employed by LISREL and related programs may produce parameter estimates that are not compatible with the measurement model in Equation 2. For instance, negative uniqueness components θ_i or factor correlations in excess of 1.0 are not uncommon. In formal terms, all estimated covariance matrices (here: Φ, Θ) are conceived as Gramian and must be non-negative definite. Violations would imply a complex-valued measurement space. Substantive considerations may produce even stricter criteria when, for instance, the communality of a measure exceeds its known reliability. For all practical purposes, the emergence of non-admissible parameter estimates indicates poor specification of the structural model.

Model fit—fit is evaluated in terms of deviations between the sample and the estimated model covariance matrices. Several measures are conceivable and have been proposed in various papers. A popular fit statistic with powerful large-sample characteristics is the maximum-likelihood G^2, computed as

$$G^2 = N\left[\ln|\hat{\Sigma}| - \ln|\mathrm{S}| + \mathrm{trace}(\mathrm{S}\hat{\Sigma}^{-1}) - p\right]. \qquad (6)$$

Under multinormality, and when the correct structural model is selected, G^2 is asymptotically χ^2-distributed with

$$df = \frac{p(p+1)}{2} - t, \tag{7}$$

where p is the order of the covariance matrix and t is the number of independently estimated model parameters. When a more restricted or modified model is applied, G^2 will follow a non-central χ^2-distribution.

Model fit can only be properly assessed when the estimation has converged to an admissible solution. Generally, inadmissible parameter estimates are associated with over-fit of the model so that the G^2-statistic will be negatively biased. It is also well-known that non-normality and non-random sampling will bias the G^2-statistic in the positive direction.

Using these four criteria (identification, convergence, admissibility, and fit), performance of congeneric trait-only factor analysis was evaluated with 23 empirical multitrait-multimethod matrices. The datasets were obtained from publications in psychological, sociological, educational and marketing research journals comprising a probably typical collection of MTMM matrices from these fields. The sample of the datasets was biased: many MTMM matrices with very small sample size, incomplete measurement design, and/or correlations based on pairwise deletion or other non-Gramian procedures were rejected. Conversely, datasets that had been reanalyzed in the literature had a higher chance to be included in the sample. Origin and nature of the matrices are described in the Appendix.

Results of the analysis are summarized in Table 1. No globally under-identified solutions of the congeneric trait-only model were observed—all the information matrices were of full rank. Convergence problems occurred with the three datasets "Attitudes to Authority (Burwen & Campbell)", "Personality Traits (Kelley & Krey)", and "Job Behavior (Dickinson & Tice)". Inspection of intermediate solutions for these datasets suggested local under-identification as the likely reason for non-convergence. In addition to three non-converged solutions, the model produced inadmissible parameter estimates for seven further datasets. This left 13 of the 23 datasets with admissible congeneric solutions.

The G^2-statistic shows acceptable fit for just two of the remaining 13 datasets ["Smoking and Capital Punishment (Jaccard)" and "Three Attitudes (Flamer, Sample 1)"]; Jaccard's dataset should not be given much weight, however, considering the small N of 35.

Table 2 shows the parameter estimates for the Flamer (Sample 1) dataset. The traits are "Attitude towards Discipline in Children" (ADC), "Attitude towards Mathematics" (AM), and "Attitude towards the Law" (AL). All assessment methods are paper-and-pencil, but comprise different item types and response formats: dichotomous Likert scales, Thurstone scales, and the semantic differential (SD) technique. Since the maximum-likelihood G^2-statistic indicates

Table 1: Congeneric trait-only analysis

Dataset	Identified	Converged	Admissible Solution	G^2	df	N
Intelligence and Effort (Mayo)			no	2.4	1	166
Intelligence & Alertness (Thorndike)				66.4	8	750
Popularity & Expansiveness (Borgatta)				96.3	19	125
Smoking and CP (Jaccard)				15.9	19	35
Leadership Study (Summers, et al.)				185.1	19	290
Authority (Burwen & Campbell)		no	no	22.5	6	57
Drives in Rats (Anderson)			no	40.9	6	50
Involvement Components (Arora)				107.6	24	96
Job Behavior (Dickinson & Tice)		no	no	144.6	24	149
Three Attitudes (Flamer, Sample 1)				23.3	24	105
Three Attitudes (Flamer, Sample 2)			no	36.0	24	105
Stress Measures (Karst & Most)			no	198.9	24	80

Table 1—*Continued*

Dataset	Identified	Converged	Admissible Solution	G^2	df	N
Job Performance (Lawler)				100.9	24	113
Moral Dilemma (Shepherd)				336.3	24	487
Contraceptives (Kothandapani)			no	369.8	51	100
Attitudes to the Church (Ostrom)				135.5	51	189
Drug Use Reports (Stacy et al.)				368.8	51	190
Clinical Clerkships (Boodoo)				287.2	87	136
Personality Traits (Kelley & Krey)		no	no	123.4	14	311
Desirability (Jackson & Singer)				1194.0	164	480
Interaction Process Vars. (Borgatta)			no	279.1	80	125
Guilford-Martin Fact. (Carroll)			no	288.2	80	110
Assessment (Kelly & Fiske)				140.5	80	124

Table 2: Congeneric trait-only estimates of the Flamer (sample 1) data

Factor loading matrix $\hat{\Lambda}_\tau$

Method	Trait	Trait factors ADC	AM	AL	Uniqueness Estimates θ
Likert	ADC	.85	.00	.00	.28
	AM	.00	.77	.00	.41
	AL	.00	.00	.61	.63
Thurstone	ADC	.84	.00	.00	.29
	AM	.00	.80	.00	.36
	AL	.00	.00	.62	.62
SD	ADC	.50	.00	.00	.75
	AM	.00	.94	.00	.12
	AL	.00	.00	.71	.50

Factor correlations $\hat{\Phi}_\tau$

	ADC	AM	AL
ADC	1.0		
AM	-.07	1.0	
AL	.39	-.05	1.0

$G^2 = 23.28$ $P = 0.503$

$df = 24$ $N = 105$

an acceptable model fit, *convergent* validity is confirmed for all nine trait-method combinations. One can also see that the relative precision of measurement is consistently higher for measures of "Attitude towards Mathematics" than for "Attitude towards the Law". Some caution is indicated for the trait "Attitude towards Discipline in Children" which shows a larger variation in the size of its factor loadings. The heterogeneity of the factor loadings is marginally significant: tested against a model with equal factor loading for each trait, there is a fit increase of Diff-$G^2 = 15.1 (df = 6, P = 0.020)$ for the congeneric model.

Discriminant validity between the trait concepts can be judged from the estimated factor correlation matrix Φ_τ. Evidently "Attitude towards Mathematics" is virtually unrelated to the other two traits, while the disattenuated correlation coefficient of $\hat{\phi}_{\tau_3,\tau_1} = .39$ shows a mild association between the attitudes towards the law and towards discipline.

In summary: Estimation of the congeneric trait-only factor model will often converge to an admissible solution, but the fit to empirical multitrait-multimethod matrices tends to be poor. These problems of model misfit can be blamed on the data—most empirical datasets did not support the notions of convergent and discriminant validity. In the two cases where the model showed a good fit, the assessment methods appeared to differ only in relatively minor aspects of question wording.

2.2 Trait-method factor analysis

As a less restrictive alternative to the trait-only model, several authors have suggested to include additional method factors (Althauser, 1974; Althauser & Heberlein, 1970; Althauser, Heberlein, & Scott, 1971; Jöreskog, 1971; Kalleberg & Kluegel, 1975; Schmitt, 1978; Werts, Jöreskog, & Linn, 1972; Werts & Linn, 1970; Werts, Linn & Jöreskog, 1971). The rationale for added method factors is that, apart from expressing trait variation, measures may also be correlated because they share the same assessment method. Method factors will purportedly account for systematic variation due to these shared method components.

For a 3-trait-by-3-method measurement design, the factor loading matrix $\Lambda_{\tau,\mu}$ of the trait-method factor model is simply constructed by augmenting Λ_τ from Equation 4 with three additional method factors μ_1, μ_2, and μ_3:

$$\Lambda_{\tau\mu} = \left(\begin{array}{ccc|ccc} \lambda_{1,\tau_1} & 0 & 0 & \lambda_{1,\mu_1} & 0 & 0 \\ 0 & \lambda_{2,\tau_2} & 0 & \lambda_{2,\mu_1} & 0 & 0 \\ 0 & 0 & \lambda_{3,\tau_3} & \lambda_{3,\mu_1} & 0 & 0 \\ \lambda_{4,\tau_1} & 0 & 0 & 0 & \lambda_{4,\mu_2} & 0 \\ 0 & \lambda_{5,\tau_2} & 0 & 0 & \lambda_{5,\mu_2} & 0 \\ 0 & 0 & \lambda_{6,\tau_3} & 0 & \lambda_{6,\mu_2} & 0 \\ \lambda_{7,\tau_1} & 0 & 0 & 0 & 0 & \lambda_{7,\mu_3} \\ 0 & \lambda_{8,\tau_2} & 0 & 0 & 0 & \lambda_{8,\mu_3} \\ 0 & 0 & \lambda_{9,\tau_3} & 0 & 0 & \lambda_{9,\mu_3} \end{array}\right) \quad (8)$$

This expanded model seems to be at least partially motivated by hopes of finding a statistical procrustes method able to eliminate method effects from the measurements. The argument goes as follows:

> If trait factors are uncorrelated with method factors, the respective factor scores should also be uncorrelated in the population. The trait scores would be retained for further analysis of "method-free" trait measures, while method scores would be rejected as "trait-less" measurement artifacts. In this sense the trait-method factor model can possibly be used to separate trait and method components.

Even though this proposition sounds somewhat fantastic, the confirmatory trait-method factor model can easily be restricted to independence between traits and methods specifying the factor intercorrelation matrix as

$$\Phi_r = \left(\begin{array}{ccc|ccc} 1.0 & \phi_{\tau_1,\tau_2} & \phi_{\tau_1,\tau_3} & 0 & 0 & 0 \\ \phi_{\tau_2,\tau_1} & 1.0 & \phi_{\tau_2,\tau_3} & 0 & 0 & 0 \\ \phi_{\tau_3,\tau_1} & \phi_{\tau_3,\tau_2} & 1.0 & 0 & 0 & 0 \\ \hline 0 & 0 & 0 & 1.0 & \phi_{\mu_1,\mu_2} & \phi_{\mu_1,\mu_3} \\ 0 & 0 & 0 & \phi_{\mu_2,\mu_1} & 1.0 & \phi_{\mu_2,\mu_3} \\ 0 & 0 & 0 & \phi_{\mu_3,\mu_1} & \phi_{\mu_3,\mu_2} & 1.0 \end{array}\right). \qquad (9)$$

The remainder of this section will examine model performance with a block-diagonal type of correlation structure described in Equation 9. When clarity is required, explicit reference is made to the *trait-method independence* models.

The difference between trait-method and trait-only factor models is not just a matter of quantity of factors. The trait-only model, conceived in the psychometric tradition of parallel measurement of a single latent trait, describes non-overlapping factor concepts. The trait-method factor model, on the other hand, expresses the systematic variance of each measure as the linear combination of *two* latent factors and is *overlapping*. In consequence, the trait-method model does not reflect the parallelity concepts of classical test theory and is largely irrelevant to the assessment of factorial validities in a set of measures. The trait-method model rather describes a metric linear decomposition of an observed measurement structure distantly related to decomposition models approaches in the tradition of Beals et al. (1968). Furthermore, the model appears to deviate substantially from the Campbell & Fiske (1959) trait conception—the validity criteria of the original paper fail to establish reasonable boundaries of the parameter space (Althauser, 1974; Althauser & Heberlein, 1970; Althauser, Heberlein, & Scott, 1971).

Performance of the trait-method factor model is summarized in Table 3. The model was globally unidentified with four datasets ["Intelligence and Effort (Mayo)", "Intelligence and Alertness (Thorndike)", "Authority (Burwen & Campbell)", and "Drives in Rats (Anderson)"]. The correlation matrices of these datasets were too small, containing fewer empirical correlation coefficients

Table 3: Trait-method independence factor analysis

Dataset	Identified	Converged	Admissible Solution	G^2	df	N
Intelligence and Effort (Mayo)	no					
Intelligence & Alertness (Thorndike)	no					
Popularity & Expansiveness (Borgatta)		no	no	3.7	5	125
Smoking and CP (Jaccard)	no: the 23rd (of 31) parameter may not be identified (ϕ_{μ_4,μ_3})					
Leadership Study (Summers, et al.)	no: the 21st (of 31) parameter may not be identified (ϕ_{μ_4,μ_1})					
Authority (Burwen & Campbell)	no					
Drives in Rats (Anderson)	no					
Involvement Components (Arora)		diverged	no	403.9	12	96
Job Behavior (Dickinson & Tice)	no: the 31st parameter (of 33) may not be identified (θ_7)	diverged				
Three Attitudes (Flamer, Sample 1)	no: the 6th parameter (of 33) may not be identified (λ_{3,μ_1})	117+				
Three Attitudes, Flamer, Sample 2)	no: the 6th parameter (of 33) may not be identified (λ_{3,μ_1})	145+				
Stress Measures (Karst & Most)		diverged	no	3892.2	12	80

Table 3—*Continued*

Dataset	Identified	Converged	Admissible Solution	G^2	df	N
Job Performance (Lawler)		diverged	no	5580.6	12	113
Moral Dilemma (Shepherd)		diverged	no	21688.9	12	487
Contraceptives (Kothandapani)			no	53.1	33	100
Attitudes to the Church (Ostrom)			no	21.7	33	189
Drug Use Reports (Stacy et al.)			no	93.5	33	190
Clinical Clerkships (Boodoo)	no: the 2nd (of 58) parameter may not be identified (λ_{1,μ_1})	diverged				
Personality Traits (Kelley & Krey)	no: the 23rd (of 31) parameter may not be identified (ϕ_{μ_2,μ_1})					
Desirability (Jackson & Singer)		no	no	410.2	134	480
Interaction Process Vars. (Borgatta)		no	no	100.5	62	125
Guilford-Martin Fact. (Carroll)			no	112.5	62	110
Assessment (Kelly & Fiske)			no	57.5	62	124

than there were parameters to be estimated. Seven further cases also resulted in unidentified solutions even though the number of correlation coefficients exceeded the number of independent model parameters ["Smoking and Capital Punishment (Jaccard)", "Leadership Study (Summers et al.)", "Job Behavior (Dickinson & Tice)", "Three Attitudes (Flamer, Sample 1)", "Three Attitudes (Flamer, Sample 2)", "Clinical Clerkships (Boodoo)", and "Personality Traits (Kelley & Krey)"]. The parameters involved vary even among similar structured datasets so that one might expect empirical underidentification. Yet, the frequency of unidentified solutions appears suspiciously high.

Among the remaining twelve datasets, solutions converged in five cases. Convergence failed in seven cases. Yet, in no case, converged or not, was the solution admissible. Table 4 shows the inadmissible solution for the Assessment (Kelley & Fiske) data as a typical example. Traditional interpretation of this well-known dataset has occasionally concluded that the Staff and Self Rating method factors should be combined because their correlation is so excessive (Jöreskog, 1971; Browne, 1984a). Such a decision supposes that the estimates are inadmissible as a consequence of problematic sample correlation matrix rather than because of a structural deficiency of the model. This does not seem to be the case. First, if inadmissibility was due to sample problems, one should also be able to find datasets that have an admissible solution. The search for such a dataset was negative, as documented in Table 3. Second, a strong point can be made that the trait-method model is conceptually flawed. Suppose that all measures in a particular study share some common variance due to any kind of shared circumstances. Common variance can be shared for a number of reasons, for instance (1) choice of similar measurement situations, (2) choice of similar traits, or (3) a strong general factor of individual differences. The three interpretations relate the common variance to method, trait, or neutral concepts, respectively, but they cannot possibly be distinguished on empirical grounds in a single MTMM study.

This conceptual identification problem has a direct numerical equivalent. The common variance may be accounted for by either the covariance structure due to trait factors or by the method factor structure. Existence of the indeterminacy can easily be demonstrated for two more restricted forms of trait-method factor analysis.

Two-factor model: Suppose the factor correlation matrix is restricted so (a) that all traits are perfectly correlated with each other and (b) all methods are likewise correlated with unity among themselves. Such a model is equivalent to an exploratory factor analytic solution with two orthogonal factors and the loading matrix

$$\Lambda_{\tau\mu} = \left(\begin{array}{c|c} \lambda_{1,\tau} & \lambda_{1,\mu} \\ \lambda_{2,\tau} & \lambda_{2,\mu} \\ \vdots & \vdots \\ \lambda_{p,tau} & \lambda_{p,\mu} \end{array} \right) \qquad (10)$$

Table 4: Trait-method independence factor analysist of the Kelly & Fiske assessment data

Factor loading matrix $\hat{\Lambda}_{\tau\mu}$

Method	Trait	Trait factors A	C	S	P	I	Method factors Staff	Mate	Self	Uniqueness Estimates θ
Staff Ratings	A	.86	.00	.00	.00	.00	-.07	.00	.00	.26
	C	.00	.83	.00	.00	.00	-.05	.00	.00	.31
	S	.00	.00	.60	.00	.00	.09	.00	.00	.62
	P	.00	.00	.00	.89	.00	.14	.00	.00	.20
	I	.00	.00	.00	.00	.72	.15	.00	.00	.45
Teammate Ratings	A	.84	.00	.00	.00	.00	.00	.13	.00	.29
	C	.00	.83	.00	.00	.00	.00	.28	.00	.47
	S	.00	.00	.68	.00	.00	.00	.35	.00	.41
	P	.00	.00	.00	.18	.00	.00	.58	.00	.65
	I	.00	.00	.00	.00	.57	.00	.50	.00	.43
Self Ratings	A	.56	.00	.00	.00	.00	.00	.00	.16	.68
	C	.00	.45	.00	.00	.00	.00	.00	.24	.76
	S	.00	.00	.44	.00	.00	.00	.00	.28	.74
	P	.00	.00	.00	.43	.00	.00	.00	.41	.66
	I	.00	.00	.00	.00	.67	.00	.00	.57	.28

Factor correlations $\hat{\Phi}_{\tau}$

	A	C	S	P	I	Staff	Mate	Self
A	1.00							
C	.56	1.00						
S	-.39	-.43	1.00					
P	.33	.62	-.07	1.00				
I	.54	.30	-.03	.46	1.00			
Staff	.00	.00	.00	.00	.00	1.00		
Mate	.00	.00	.00	.00	.00	.88	1.00	
Self	.00	.00	.00	.00	.00	-2.34	-.01	1.0

$G^2 = 57.64$ $P = 0.503$

$df = 62$ $N = 105$

It is well known that only $2p - 1$ of the $2p$ parameters in matrix 10 can be estimated, one parameter has to be set to zero in order to fix the orientation of the factors (Anderson & Rubin, 1956; Dunn, 1973; Jöreskog & Sörbom, 1979, pp. 40–43). The trait-method factor model does not incorporate such a constraint and is therefore unidentified.

Equally weighted indicators: Suppose the factor model is simplified so that all estimated loadings of a given factor have the same value. Such a solution, presented in equation 11, specifies that variance of a given factor is equally reflected in all its indicators. This would be generally attractive and simplify greatly the interpretation of the factors.

$$\Lambda_{\tau\mu} = \left(\begin{array}{ccc|ccc} \lambda_1 & 0 & 0 & \lambda_4 & 0 & 0 \\ 0 & \lambda_2 & 0 & \lambda_4 & 0 & 0 \\ 0 & 0 & \lambda_3 & \lambda_4 & 0 & 0 \\ \lambda_1 & 0 & 0 & 0 & \lambda_5 & 0 \\ 0 & \lambda_2 & 0 & 0 & \lambda_5 & 0 \\ 0 & 0 & \lambda_3 & 0 & \lambda_5 & 0 \\ \lambda_1 & 0 & 0 & 0 & 0 & \lambda_6 \\ 0 & \lambda_2 & 0 & 0 & 0 & \lambda_6 \\ 0 & 0 & \lambda_3 & 0 & 0 & \lambda_6 \end{array}\right) \tag{11}$$

$\Lambda_{\tau\mu}$ can be expressed as the product of a design matrix $\mathbf{A}_{\tau\mu}$ and a diagonal matrix $\mathbf{D}_{\tau\mu}$ of factor loadings

$$\begin{aligned} \Lambda_{\tau\mu} &= \left(\begin{array}{ccc|ccc} 1 & 0 & 0 & 1 & 0 & 0 \\ 0 & 1 & 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 1 & 0 & 0 \\ 1 & 0 & 0 & 0 & 1 & 0 \\ 0 & 1 & 0 & 0 & 1 & 0 \\ 0 & 0 & 1 & 0 & 1 & 0 \\ 1 & 0 & 0 & 0 & 0 & 1 \\ 0 & 1 & 0 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 & 0 & 1 \end{array}\right) \left(\begin{array}{cccccc} \lambda_1 & 0 & 0 & 0 & 0 & 0 \\ 0 & \lambda_2 & 0 & 0 & 0 & 0 \\ 0 & 0 & \lambda_3 & 0 & 0 & 0 \\ \hline 0 & 0 & 0 & \lambda_4 & 0 & 0 \\ 0 & 0 & 0 & 0 & \lambda_5 & 0 \\ 0 & 0 & 0 & 0 & 0 & \lambda_6 \end{array}\right) \\ &= \qquad \mathbf{A}_{\tau\mu} \qquad\qquad \mathbf{D}_{\lambda} \end{aligned} \tag{12}$$

The model equation then becomes

$$\Sigma_x = \mathbf{A}_{\tau\mu}\mathbf{D}_{\lambda}\Phi_{\tau\mu}\mathbf{D}_{\lambda}\mathbf{A}'_{\tau\mu} + \Theta \tag{13}$$

Since $\mathbf{A}$ has rank $t + m - 1$, only at most $t + m - 1$ functions of the $t + m$ factors are estimable (Graybill, 1961, pp. 228–229). Wothke (1984) has shown that any single parameter in Equation 13 can be fixed in a way that solves the identification problem. Such a solution would, however, be arbitrary and render the remaining parameter estimates meaningless.

Discussion of identification conditions for the general form of the trait-method model is still difficult and far from conclusive. It can, for instance, be shown that the trait-method independence model does not fulfill the sufficiency conditions for factor identification outlined by Anderson & Rubin (1956), Jennrich (1978), and Jöreskog & Sörbom (1979). According two these sufficiency conditions, we may have a case of rotational underidentification on our hands, but there is no conclusive proof. For the moment we shall be satisfied that the case of *equally weighted indicators*, a textbook example of a simple structure decomposition, cannot be identified. The relative orientation of the trait and method subspaces remains undefined.

2.3 Discussion

Twentythree empirical MTMM matrices were analyzed with trait-only and trait-method independence factor models. Neither model showed perfect performance with all datasets. The *trait-only* model converged to admissible and identified solutions in more than half the cases, but model fit was acceptable only in two cases. Trait-only factor analysis is the most desirable model, but most empirical correlation matrices do not conform.

On the other hand, analyses with the *trait-method independence* model failed completely. The practical consequence of these results is that the trait-method model is not applicable to any of the 23 datasets. Apparently, the factor analytic treatment of the multitrait-multimethod matrix has reached its limits with the trait-method model already. The seemingly sensible approach of reducing the systematic variance into sets of trait and method factors cannot be applied. The solutions are either not identified or are not admissible. Either case precludes substantive interpretation of the parameter estimates. The reason is that the structural conception of the measurement design is deficient. The trait-method model appears to be overparameterized with the consequence that the solutions are rotationally undetermined.

3 Covariance Component Analysis

Covariance component analysis (CCA) was first introduced by Bock (1960) and Bock & Bargmann (1966) as a multivariate random model for factorial measurement designs. The method was originally designated as "covariance structure analysis", the term is avoided here because it has since become synonymous with the more general class of structural equation models. A successful application by Bock, Dicken, & Van Pelt (1969) investigates the effects of content-acquiescence interaction in MMPI scales.

Covariance component analysis explicitly accounts for the general level of covariation common to all measures in the design, trait variation, and plus method variation, but contains only those parameters that, at least in principle, can be estimated. CCA thus avoids the indeterminacy encountered with trait-method factor analysis. Apparent problems with CCA were rooted in traditionally strict assumptions about scale and error variance of each measure, untenable for MTMM correlation matrices. This section introduces a terminology of generalized covariance component models appropriate for scale free analysis. Discussion is restricted to structural characteristics of CCA; parameter estimates are always obtained with the LISREL (Jöreskog & Sörbom, 1986) program.

3.1 Covariance component structures

In the original formulation by Bock & Bargmann (1966), covariance component analysis describes the facet-structured observed variables as linear functions of underlying latent variates. The set of measures shows the latent structure decomposition

$$\mathbf{X}_{(n\times mt)} = \Xi_{n\times(1+t+m)}\mathbf{A}'_{(1+t+m)\times mt} + \mathbf{E}_{(n\times mt)} \qquad (14)$$

The matrix Ξ shows one variate for general latent variation, t variates for the traits, and m variates for the methods. The structural coefficient matrix is assumed to be fully known and, for an MTMM matrix with 3 traits and 3 methods, takes the form

$$\mathbf{A} = \begin{pmatrix} 1 & 1 & 0 & 0 & 1 & 0 & 0 \\ 1 & 0 & 1 & 0 & 1 & 0 & 0 \\ 1 & 0 & 0 & 1 & 1 & 0 & 0 \\ 1 & 1 & 0 & 0 & 0 & 1 & 0 \\ 1 & 0 & 1 & 0 & 0 & 1 & 0 \\ 1 & 0 & 0 & 1 & 0 & 1 & 0 \\ 1 & 1 & 0 & 0 & 0 & 0 & 1 \\ 1 & 0 & 1 & 0 & 0 & 0 & 1 \\ 1 & 0 & 0 & 1 & 0 & 0 & 1 \end{pmatrix} \qquad (15)$$

g traits methods

The correspondence to Equation 8 is apparent. The expectation of the sample covariance matrix $\mathbf{S}_x = \frac{1}{n-1}\mathbf{X}'(\mathbf{I} - \frac{1}{n}\mathbf{1}\mathbf{1}')\mathbf{X}$ is

$$\Sigma_x = \mathcal{E}(\mathbf{S}_x) = \mathbf{A}\Phi\mathbf{A}' + \Theta, \qquad (16)$$

with Φ being the covariance matrix of the latent variates and Θ the (typically diagonal) covariance matrix of unique and error components. Error variances (the diagonal entries in Θ) may or may not be restricted to homoscedasticity. Bock & Bargmann originally assumed that Φ is diagonal. This assumption is unnecessarily strict for applied purposes, however, and Wiley et al. (1973) and Jöreskog (1978) have extended the model to include correlated latent structures.

Graybill (1961, pp. 228–229) has shown that, since $\mathbf{A}$ in Equation 16 is not of full rank, not all parameters in Φ and Θ can be estimated. However, estimation of the essential variance components can be attained via reparameterization. To this effect, two matrices $\mathbf{K}$ and $\mathbf{L}$ with

$$\mathbf{A} = \mathbf{K} \cdot \mathbf{L}, \tag{17}$$

are chosen, so that $\mathbf{K}$ is a matrix of $m+t-1$ orthonormal column contrasts in $\mathbf{A}$ and $\mathbf{L}$ is a matrix of $m+t-1$ orthogonal row contrasts of $\mathbf{A}$ with

$$\mathbf{L} = (\mathbf{K}'\mathbf{K})^{-1}\mathbf{K}'\mathbf{A} = \mathbf{K}'\mathbf{A}. \tag{18}$$

$\mathbf{X}$ may then be a expressed in terms of a reduced $n \times (t+m-1)$ latent structure matrix $\Xi^* = \Xi \cdot \mathbf{L}'$ as

$$\mathbf{X}_{(n\times mt)} = \Xi^*\mathbf{K}' + \mathbf{E} \tag{19}$$

$$= \Xi\mathbf{L}'\mathbf{K}' + \mathbf{E}. \tag{20}$$

Instead of Φ, the covariance matrix of the latent components, now $\Phi^* = \mathbf{L}\Phi\mathbf{L}'$, the covariance matrix of orthogonal transforms of the original latent components, is estimated. The matrix Φ^* has two fewer rows and columns than Φ, but, since the omitted parameters could not be estimated in the first place, no information is effectively lost. Interpretation must be based on the transformed parameters in Φ^* which correspond to the three groups of variates:

one variate for the general level of covariation,

$t-1$ variates describing differences in covariation due to traits, and

$m-1$ variates expressing differences due to methods.

The reparameterization transforms Equation 16 into

$$\Sigma_x = \mathcal{E}(\mathbf{S}_x) = \mathbf{K}\Phi^*\mathbf{K}' + \Theta, \tag{21}$$

Interpretation of the covariance components in Φ^* must reflect the particular choice of contrasts in $\mathbf{K}$ in addition to the empirical covariance structure. In the case of a 3 traits by 3 methods measurement design, for instance, $\mathbf{K}$ may be chosen as

$$\mathbf{K} = \begin{pmatrix} 1/3 & \sqrt{2}/3 & 0 & \sqrt{2}/3 & 0 \\ 1/3 & -1/\sqrt{18} & 1/\sqrt{6} & \sqrt{2}/3 & 0 \\ 1/3 & -1/\sqrt{18} & -1/\sqrt{6} & \sqrt{2}/3 & 0 \\ 1/3 & \sqrt{2}/3 & 0 & -1/\sqrt{18} & 1/\sqrt{6} \\ 1/3 & -1/\sqrt{18} & 1/\sqrt{6} & -1/\sqrt{18} & 1/\sqrt{6} \\ 1/3 & -1/\sqrt{18} & -1/\sqrt{6} & -1/\sqrt{18} & 1/\sqrt{6} \\ 1/3 & \sqrt{2}/3 & 0 & -1/\sqrt{18} & -1/\sqrt{6} \\ 1/3 & -1/\sqrt{18} & 1/\sqrt{6} & -1/\sqrt{18} & -1/\sqrt{6} \\ 1/3 & -1/\sqrt{18} & -1/\sqrt{6} & -1/\sqrt{18} & -1/\sqrt{6} \end{pmatrix}. \tag{22}$$

g — traits — methods

Then the first *trait* contrast will reflect individual differences between trait 1 and the other two, the second one expresses difference variation between traits 2 and 3. When the variance due to these contrasts is zero, the original traits are indistinguishable (i.e., perfectly correlated). Traits can only be distinguished when the contrast variates show non-zero variance. *Method* variates would be interpreted correspondingly as differential responses to assessment methods.

When substantive considerations permit, the contrast matrix **K** should be simplified at one of the following levels: One, column contrasts for general, trait, and method components may be chosen as blockwise orthogonal. This allows testing independence between the trait and method differences and the general variate. Two, if *all* columns in **K** are orthogonal, correlations derived from Φ^* may be directly interpreted, but the variance estimates will still be functionally dependent on the scale of the contrasts. Finally, when all contrasts are orthonormal (i.e., orthogonal *and* normalized to unit length), all parameters in Φ^* are estimated on the same scale and latent variances can be compared relative to each other. Orthonormal contrasts are advantageous when the relative contribution of trait or method facets is assessed. All data analyses in this paper are based on orthonormal contrast matrices.

Several types of covariance component models, defined by restrictions of the matrix Φ^*, should be distinguished.

Fully correlated Φ^*: The observed covariance matrix can be expressed as a compound of trait and method variance components:

$$\Phi^* = \left(\begin{array}{c|c|c} \sigma_2^{*2} & \text{(symm.)} & \\ \hline \sigma^*_{g\tau} & \Phi^*_{\tau\tau} & \\ \hline \sigma^*_{g\mu} & \Phi^*_{\mu\tau} & \Phi^*_{\mu\mu} \end{array}\right) \tag{23}$$

There is some justification for trait concepts, but general, trait, and method variates are correlated.

Independent-common-variation: The first row and column show zero entries in the off-diagonal elements:

$$\Phi^* = \left(\begin{array}{c|c|c} \sigma_g^{*2} & \text{(symm.)} & \\ \hline 0 & \Phi^*_{\tau\tau} & \\ \hline 0 & \Phi^*_{\mu\tau} & \Phi^*_{\mu\mu} \end{array}\right) \tag{24}$$

Trait and method variation is independent of the general variate, but trait and method contrasts may be intercorrelated.

Trait-method independence: Trait contrasts are independent of method contrasts ($\Phi^*_{\mu\tau} = 0$), but the general factor may covary with either.

$$\Phi^* = \left(\begin{array}{c|c|c} \sigma_g^{*2} & \text{(symm.)} & \\ \hline \sigma^*_{g\tau} & \Phi^*_{\tau\tau} & \\ \hline \sigma^*_{g\mu} & 0 & \Phi^*_{\mu\mu} \end{array}\right) \tag{25}$$

Block-diagonal Φ^*: Trait contrasts are uncorrelated with method contrasts and, in addition, the general variate is independent of both trait and method contrasts.

$$\Phi^* = \left(\begin{array}{c|c|c} \sigma_g^{*2} & \text{(symm.)} & \\ \hline 0 & \Phi^*_{\tau\tau} & \\ \hline 0 & 0 & \Phi^*_{\mu\mu} \end{array}\right) \tag{26}$$

If the empirical measurement structure has a block-diagonal covariance component form, then the following three conclusions are legitimate: (a) patterns of individual differences in traits do not predict individual differences in response to methods; (b) differential response to method does not predict an individual's average level on all measures; and (c) trait contrasts do not predict the individual's relative standing on the general variate.

Diagonal Φ^*: All (reparameterized) variates are uncorrelated. Diagonality is postulated by design only in the case of 2^n measurement designs. When trait or method facets contain more than two elements, diagonality will partly depend on the particular choice of contrasts. In these cases, contrast selection must be guided by substantive theory. Diagonality implies that the researcher has, in substantive terms, found a most parsimonious account of the observed covariance structure. This transcends the question whether trait and method differences are independent.

3.2 Covariance component analysis with unknown scale factors

Fixed-scale CCA in the form of Equation 21 calls for known or hypothesized scales of the latent variates over the entire set of measures or, alternatively, necessitates specific assumptions about the uniqueness components in the diagonal of Θ. Knowledge of the scale of measurement, however, is often not available and, just as frequently, is of secondary interest in analytic behavioral research. For instance, scale information in correlation analysis is lost entirely due to standardization. Furthermore, multitrait-multimethod analyses are frequently conducted when fixed-scale assumptions across different traits and methods are not meaningful on conceptual grounds. Standardization of observed variables in these cases imposes an arbitrary ceiling on the variance of the observed variables and, in order to obtain any kind of interpretable estimates for Φ^* in the linear model framework, the relative true score scales have to be estimated. Wiley, Schmidt & Bramble (1973) propose a class of scale-free generalizations of Equation 21 by introducing a diagonal matrix G of scaling constants:

$$\Sigma_x = \mathbf{G}\mathbf{K}\Phi^*\mathbf{K}'\mathbf{G} + \Theta. \tag{27}$$

G will absorb scale differences among the observed measures and should be interpreted accordingly. Wiley et al. (p. 317) state that

> The major utility of ... [G] is for dealing with those situations in which the observed variables are measured in different metrics. For such cases the introduction of ... [G] whose elements do not have to be related to the variances of the variables allows for optimal rescaling.

3.2.1 Model identification

One element in G and Φ^* must be set to a positive value to fix the scale of the estimates. This is because the Gramian product in Equation 27 is generally not identified due to a scale trade-off between Φ^* and G: multiplication of Φ^* with a positive constant a is fully compensated for when G is simultaneously divided by $\sqrt{a}$. This trivial underidentification has no consequence for the substantive *interpretation* of the parameter estimates. Only the *relative* size of the component variance and covariance estimates in Φ^* is required to reconstruct the latent correlation structure of the measures. Furthermore, estimates of the scaling constants in G are only meaningful in conjunction with the estimate of Φ^*.

The underidentification is removed by a single non-zero constraint; all exemplary analyses will use the identity

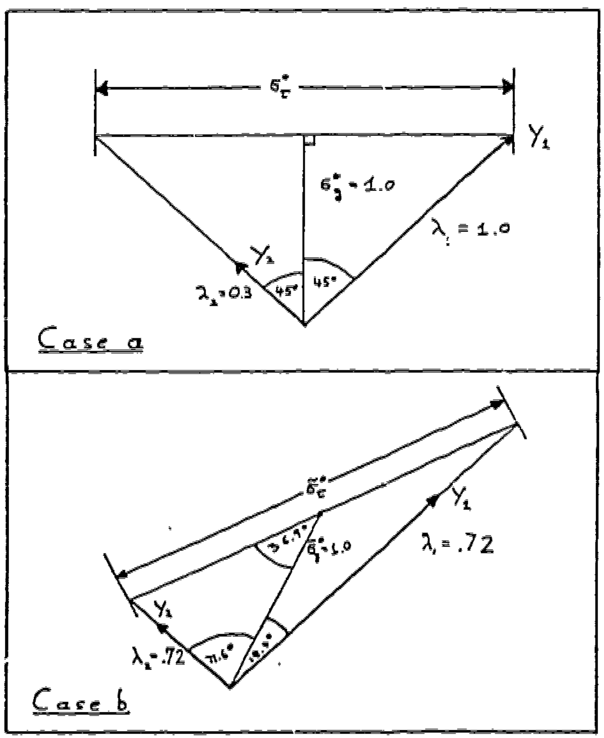
$$\sigma_g^{*2} = 1, \tag{28}$$

restricting the variance of the general variate to unity. Variance estimates for the trait and method contrasts have to be evaluated relative to the variance of the general variate.

An additional and more complicated identification problem arises with *fully correlated* and *trait-method independence* CCA. It turns out that correlations with the general variate; i.e., the elements of $\sigma_{g\tau}^*$ and $\sigma_{g\mu}^*$ are unidentified in the scale-free model when the measurement design is small. Figure 1 illustrates this identification problem with a simple two-dimensional case. In both parts of the Figure, the length of the general variate g has been fixed to unity ($\sigma_g^{*2} = \tilde{\sigma}_g^{*2} = 1.0$). Case a describes the latent measurement structure as diagonal or block-diagonal and with unequal scale factors ($\lambda_1 = 1.0$ and $\lambda_2 = 0.3$). Case b describes the identical measurement structure using a fully correlated version of Φ^* with $\sigma_{g\tau}^*$ corresponding to a correlation of $\cos(36.9°) = 0.80$ and equal sized scale factors ($\lambda_1 = \lambda_2 = 0.72$). Many other equivalent solutions exist and the estimation equations are undetermined.

The identification problem is practically independent of the size of the measurement design. It certainly remains when all but one contrast variates are of length zero. When the design is larger than 2 traits by 2 methods, trait-method independence CCA may be numerically identified, but some asymptotic correlations among the estimates usually exceed 0.95. Then, the precision of the parameter estimation will not be acceptable.

Figure 1: Two equivalent covariance component repr~entations



In short, scale-free diagonal, block-diagonal, and independent-common-variation CCA models are generally identified, while the scale-free versions of the fully correlated and trait-method independence models are not. These results contradict some of the claims by Wiley et al. (1973). Correlations of the common variate can be evaluated if the scale factors are known. Alternatively, scale factors can be determined when strict assumptions about the correlation structure of the common variate are imposed.

3.3 Empirical application of scale-free CCA

This section applies scale-free *block-diagonal* and *independent-common-variation* CCA to the 23 empirical MTMM matrices. Both models are direct generalizations of congeneric trait-only factor analysis. Model performance is evaluated in terms of identification, convergence, admissibility, and fit. Definitions for the first three criteria are identical to those for factor analysis in Section 2.1.

Admissibility of CCA solutions implies that Θ and Φ^* are both non-negative definite. Admissibility shall also denote that no latent variate will account for more than the total variance of the measured variables:

$$0 \leq g_i |k_{i,\tau_j}| \sigma^*_{\tau_j} \leq \sigma_i \qquad . \tag{29}$$

For correlation matrices, the upper bound becomes unity.

Models using the other three covariance component structures; i.e., strictly diagonal, trait-method independence, and fully correlated CCA are not evaluated here for different reasons. The diagonal submodel requires theory-guided, not simply design-guided, selection of contrasts and transcends the scope of this paper. The other two submodels allow for correlation between contrasts and the general variate and are ridden with identification problems.

Table 5 summarizes the scale-free block-diagonal covariance component analyses of the 23 datasets. All solutions are identified and converged, but not all are admissible. Solutions for 13 of the 23 datasets violate the admissibility conditions. Five of the 10 admissible solutions have good fit ["Popularity and Expansiveness (Borgatta)", "Smoking and Capital Punishment (Jaccard)", "Job Behavior (Dickinson & Tice)", "Job Performance (Lawler)", and "Attitudes to the Church (Ostrom)"]. With an additional dataset ["Assessment (Kelly & Fiske)"] the fit is marginal. Four of these solutions show a significant fit increase compared to the congeneric trait-only factor solution, while one ["Smoking and Capital Punishment (Jaccard)"] was already well fit by the congeneric two-trait factor model. In one case ["Job Behavior (Dickinson & Tice)"] the trait-only factor model had not converged to an admissible solution, while admissibility was achieved for the "Three Attitudes (Flamer, Sample 1)" data using the trait-only model, but not in the case of block-diagonal CCA.

The independent-common variation model is a generalization of block-diagonal structures. Allowing for non-zero covariances between trait and method

Table 5: Scale-free covariance component analysis of the datasets, block-diagonal model.

Dataset	Identified	Converged	Admissible Solution	G^2	df	N
Intelligence and Effort (Mayo)			no	0.0	0	166
Intelligence & Alertness (Thorndike)				17.2	5	750
Popularity & Expansiveness (Borgatta)				14.6	13	125
Smoking and CP (Jaccard)				9.9	13	35
Leadership Study (Summers et al.)			no	81.4	13	196
Authority (Burwen & Campbell)			no	16.8	5	57
Drives in Rats (Anderson)			no	3.5	5	50
Involvement Components (Arora)				50.4	21	96
Job Behavior (Dickinson & Tice)				16.3	21	149
Three Attitudes (Flamer, Sample 1)			no	19.3	21	105
Three Attitudes (Flamer, Sample 2)			no	35.9	21	105
Stress Measures (Karst & Most)			no	46.4	21	80
Job Performance (Lawler)				29.2	21	113

Table 5—*Continued*

Dataset	Identified	Converged	Admissible Solution	G^2	df	N
Moral Dilemma (Shepherd)				131.3	21	487
Contraceptives (Kothandapani)			no	104.5	45	100
Attitudes to the Church (Ostrom)				58.4	45	189
Drug Use Reports (Stacy et al.)			no	224.2	45	190
Clinical Clerkships (Boodoo)			no	143.1	77	136
Personality Traits (Kelley & Krey)			no	15.2	13	311
Desirability (Jackson & Singer)				848.8	154	480
Interaction Process Vars. (Borgatta)			no	211.9	77	125
Guilford-Martin Fact. (Carroll)			no	182.7	77	110
Assessment (Kelly & Fiske)				104.7	77	124

27

Table 6: Scale-free covariance component analysis of the datasets, independent-common-variation model

Dataset	Identified	Converged	Admissible Solution	G^2	df	N
Intelligence and Effort (Mayo)	not identified (the design is too small)					
Intelligence & Alertness (Thorndike)			no	4.4	3	750
Popularity & Expansiveness (Borgatta)			no	10.7	10	125
Smoking and CP (Jaccard)				8.6	10	35
Leadership Study (Summers, et al.)			no	67.0	10	196
Authority (Burwen & Campbell)	no: the first parameter (g_{11}) is not identified					
Drives in Rats (Anderson)			no	3.4	3	50
Involvement Components (Arora)				41.3	17	96
Job Behavior (Dickinson & Tice)				14.0	17	149
Three Attitudes (Flamer, Sample 1)			no	17.2	17	105
Three Attitudes (Flamer, Sample 2)			no	29.9	17	105
Stress Measures (Karst & Most)			no	43.1	17	80

Table 6—*Continued*

Dataset	Identified	Converged	Admissible Solution	G^2	df	N
Job Performance (Lawler)				22.3	17	113
Moral Dilemma (Shepherd)				94.8	17	487
Contraceptives (Kothandapani)			no	87.3	39	100
Attitudes to the Church (Ostrom)			no	52.6	39	189
Drug Use Reports (Stacy et al.)			no	207.7	39	190
Clinical Clerkships (Boodoo)			no	138.9	69	136
Personality Traits (Kelley & Krey)			no	7.7	10	311
Desirability (Jackson & Singer)				798.5	142	480
Interaction Process Vars. (Borgatta)			no	187.9	69	125
Guilford-Martin Fact. (Carroll)			no	153.2	69	110
Assessment (Kelly & Fiske)				87.4	69	124

contrasts, $(t-1) \times (m-1)$ additional parameters are estimated from the data. Results are summarized in Table 6. Solutions for two datasets are not identified: with the "Intelligence and Effort (Mayo)" data, the 2×2 measurement design is simply too small, while the "Authority (Burwen & Campbell)" correlation matrix apparently presents an empirical identification problem. Estimation procedures converged to an admissible solution in only 7 cases. Fit improvement is small in the four cases where the block-diagonal model already approximates the empirical correlation matrix ["Smoking and Capital Punishment (Jaccard)", "Job Behavior (Dickinson & Tice)", "Job Performance (Lawler)"], and "Assessment (Kelly & Fiske)". With three other datasets, for which the block-diagonal structure did not provide an acceptable account, fit improvement by the independent-common-variation model is substantial but not large enough. For the "Involvement Components (Arora)" (Diff-$G^2 = 9.1$, df $= 4$), the "Moral Dilemma (Shepherd)" (Diff-$G^2 = 36.5$, df $= 4$), and the "Desirability (Jackson & Singer)" data (Diff-$G^2 = 50.3$, df $= 12$), neither block-diagonal nor independent-common-variation covariance component structures yield close descriptions of the empirical correlation matrices.

The independent-common-variation model provides acceptable descriptions for four of the 23 empirical correlation matrices. These matrices are, however, already well accounted for by the more restricted block-diagonal covariance component model. The incremental utility of the independent-common-variation model over the block-diagonal model therefore cannot be clearly affirmed. This lack of significant improvement is quite likely a function of the particular datasets used in this study and does not indicate any model deficiency. For the time being, the model may or may not be endorsed, pending some less ambiguous evidence becoming available.

3.4 Interpretation of CCA solutions

In addition to finding the correct component structure, substantive interpretation of the parameter estimates is a necessary part of the data analysis. Unfortunately, in my experience, it is quite a difficult enterprise to explain covariance component estimates (in $\hat{\Phi}^*$) to social scientists and even to some trained statisticians. Training in contrast techniques appears to be lacking. Prior knowledge on these matters does, however, aid in the understanding of this section; Bock (1975) and Finn (1974) provide useful terminology on these matters.

As a typical example for MTMM analysis, Table 7 displays a contrast matrix **K** for the Kelly and Fiske assessment data previously discussed in Table 4 in the factor analytic context. **K** contains seven contrasts, one for the general variate, four for trait variation, and two for method variation. The matrix is columnwise orthonormal; i.e., the contrasts are uncorrelated and have unit length. The component values for the general variate (in the first column) are standard and should not be modified. Yet, since the data analysis was only concerned with determining the overall covariance component structure, and no further

Table 7: Orthonormal contrast matrix K for the Kelly & Fiske assessment data.

Method	Trait	δ_g	δ_{τ_1}	δ_{τ_2}	δ_{τ_3}	δ_{τ_4}	δ_{μ_1}	δ_{μ_2}
Staff Ratings	Assertive	.25820	.51640	0.0	0.0	0.0	.36515	0.0
	Cheerful	.25820	-.12910	.5	0.0	0.0	.36515	0.0
	Serious	.25820	-.12910	-.16667	.47141	0.0	.36515	0.0
	Poise	.25820	-.12910	-.16667	-.23570	.40825	.36515	0.0
	Interests	.25820	-.12910	-.16667	-.23570	-.40825	.36515	0.0
Team-mate Ratings	Assertive	.25820	.51640	0.0	0.0	0.0	-.18257	.31623
	Cheerful	.25820	-.12910	.5	0.0	0.0	-.18257	.31623
	Serious	.25820	-.12910	-.16667	.47141	0.0	-.18257	.31623
	Poise	.25820	-.12910	-.16667	-.23570	.40825	-.18257	.31623
	Interests	.25820	-.12910	-.16667	-.23570	-.40825	-.18257	.31623
Self Ratings	Assertive	.25820	.51640	0.0	0.0	0.0	-.18257	-.31623
	Cheerful	.25820	-.12910	.5	0.0	0.0	-.18257	-.31623
	Serious	.25820	-.12910	-.16667	.47141	0.0	-.18257	-.31623
	Poise	.25820	-.12910	-.16667	-.23570	.40825	-.18257	-.31623
	Interests	.25820	-.12910	-.16667	-.23570	-.40825	-.18257	-.31623

confirmatory substantive hypotheses were employed, the Helmert contrasts in the trait and method blocks were chosen arbitrarily. The first trait contrast determines the latent variate δ_{τ_1} as the difference between *Assertiveness* and the remaining four traits, δ_{τ_2} describes the difference between *Cheerfulness* and the average of *Seriousness*, *Unshakable Poise*, and *Broad Interests*, δ_{τ_4}, finally, contrasts *Unshakable Poise* with *Broad Interests*. Correspondingly, the method contrast δ_{μ_1} is defined to absorb the difference between *Staff Ratings* and the *average of Teammate and Self Ratings*, while δ_{μ_2} compares *Teammate Ratings* against *Self Ratings*.

Fit of the block-diagonal CCA model is marginally significant (with $G^2 =$ 104.7 and $df = 77$) and can be considered satisfactory, given that the correlations were computed from rating scales. Estimates for model Equation 27 are displayed in Tables 8 and 9. The uniqueness coefficients $(\hat{\Theta})_{ii}$ in Table 8 have the same interpretation as their factor analytic equivalent. Self ratings of *Assertiveness*, *Cheerfulness*, and *Seriousness* have uniqueness components almost twice as large as the corresponding ratings obtained from teammates and staff members, indicating that the self ratings are less reliable, reflect different insights and standards, and/or are mediated by additional constructs like the person's degree of confidence. Teammate and Staff ratings differ most noticably for *Unshakable Poise*, the unique component being twice as large for the teammate data.

The scale factor estimates $(\hat{\mathbf{G}})_{ii}$ in Table 8 reflect differences in "true score" variance of the observed measures. Measures associated with larger scale fac-

Table 8: Estimated scale factors $\hat{\mathbf{G}}$ and uniqueness coefficients $\hat{\Theta}$.

Method	Trait	Scale Factors $(\hat{\mathbf{G}})_{ii}$	Uniqueness Coeffs. $(\hat{\Theta})_{ii}$
Staff Ratings	Assertive	1.628	.250
	Cheerful	1.650	.330
	Serious	.774	.639
	Poise	2.253	.425
	Interests	1.836	.533
Team-mate Ratings	Assertive	1.616	.255
	Cheerful	1.468	.478
	Serious	.956	.404
	Poise	1.126	.837
	Interests	1.907	.418
Self Ratings	Assertive	.925	.714
	Cheerful	.987	.709
	Serious	.557	.811
	Poise	1.725	.569
	Interests	1.975	.367

tors discriminate on a relatively larger scale, above and beyond the systematic variance due to the covariance component structure $\mathbf{K}\Phi^*\mathbf{K}'$. Such an interpretation is correct for MTMM covariance matrices. When generalized CCA is based on MTMM *correlation* matrices, instead, scale factor estimates will also be dependent on the error variance of the original (unstandardized) measures. Then, the substantive interpretation of $(\hat{\mathbf{G}})_{ii}$ will be less direct. In either case, the diagonal of $(\hat{\mathbf{G}})$ contains the estimated scale factors needed to optimally rescale the original variables as $\mathbf{Y} = \mathbf{X}\mathbf{G}^{-1}$, transforming Equation 8 to

$$\begin{aligned} \Sigma_y &= \mathbf{G}^{-1}\Sigma_x\mathbf{G}^{-1} && (30) \\ &= \mathbf{G}^{-1}(\mathbf{G}\mathbf{K}\Phi^*\mathbf{K}'\mathbf{G} + \Theta)\mathbf{G}^{-1} && (31) \\ &= \mathbf{K}\Phi^*\mathbf{K}' + \mathbf{G}^{-1}\Theta\mathbf{G}^{-1} && (32) \end{aligned}$$

All three ratings of *Broad Interests* are found to show scale factors of comparable magnitude. *Assertiveness*, *Cheerfulness*, and *Seriousness* have similar scale factors for staff and teammate rating methods, while self rating factors are substantially smaller. Self ratings of these variables are not comparable to ratings made by others. For the Kelly & Fiske assessment data, scale factors and uniqueness coefficients reflect different aspects of the same phenomenon.

Entries in $\hat{\Phi}^*$ are harder to interpret than scale factors and uniqueness coefficients. Given the contrast definitions in Table 7, the estimates can be viewed

Table 9: Estimated Covariance Component Matrix $\hat{\Phi}^*$.

	δ_g	δ_{τ_1}	δ_{τ_2}	δ_{τ_3}	δ_{τ_4}	δ_{μ_1}	δ_{μ_2}
δ_g	1.0*						
δ_{τ_1}	0.0*	.670			(symmetric)		
δ_{τ_2}	0.0*	.387	.793				
δ_{τ_3}	0.0*	-.639	-.821	1.284			
δ_{τ_4}	0.0*	-.087	.107	.020	.169		
δ_{μ_1}	0.0*	0.0	0.0	0.0	0.0	.126	
δ_{μ_2}	0.0*	0.0	0.0	0.0	0.0	.100	.360

as relative variance components of the respective latent variates. The first component (ϕ^*_{11}) is fixed at unity and defines the scale of all other estimates in Φ^* and G. The remaining diagonal elements of Φ^* contain the relative variance due to the trait and method contrasts. Substantive interpretation of the variance components is only meaningful when the corresponding contrast has itself any substantive significance. For instance, $\phi^*_{\mu_1,\mu_1}$ can be understood as the variance of the difference between staff ratings versus teammate and self ratings. On the other hand, trait contrasts had been chosen arbitrarily, so that the estimates in $\Phi^*_{\tau\tau}$ have no direct substantive interpretation. While it is not inconceivable to attempt a direct interpretation of the estimates in $\Phi^*_{\tau\tau}$, the result of such an attempt would appear contrived and nearly incomprehensible.

Similar interpretative problems appear for the off-diagonal elements in Φ^*; i.e., for the covariance components that reflect the association among the contrasts variates. The moderately positive correlation between the two method variates can, for instance, be interpreted to relate all *ratings by others*: People with higher average values in staff ratings than in combined mate and self ratings also tend to be rated higher by the teammates than they rate themselves. Here again, considering the present dataset, direct interpretation of trait covariance components is not easily communicated.

When contrasts are arbitrarily selected, as in the present case, some purely exploratory transformation of the solution may be required for substantive interpretation. In the case of block-diagonal CCA, and only then, can several transformations from the tool box of the multivariate literature be reasonably employed, such as *dispersion component* comparisons, *canonical decomposition* of the covariance component blocks, and *blockwise varimax rotation*.

3.4.1 Dispersion components

The determinant of a covariance matrix is frequently regarded as a scalar measure of the generalized multivariate variance (Green & Carroll, 1976; Kendall &

Stuart, 1968; Wilks, 1932). It is well known (cf., Searle, 1982, p. 258) that the determinant of a block-diagonal matrix equals the product of the determinants of the blocks. When the covariance component solution is block-diagonal with

$$\Phi^* = \left(\begin{array}{c|c|c} \sigma_g^{*2} & \text{(symm.)} & \\ \hline 0 & \Phi_{\tau\tau}^* & \\ \hline 0 & 0 & \Phi_{\mu\mu}^* \end{array} \right), \tag{33}$$

the computation of the determinant $|\Phi^*|$ is facilitated in the form

$$|\Phi^*| = |\sigma_g^{*2}| \cdot |\Phi_{\tau\tau}^*| \cdot |\Phi_{\mu\mu}^*|. \tag{34}$$

In the case of generalized CCA, with the scale constraint of $\sigma_g^{*2} = 1$, the determinant simplifies further to

$$|\Phi^*| = |\Phi_{\tau\tau}^*| \cdot |\Phi_{\mu\mu}^*|. \tag{35}$$

The dispersion of the whole covariance component matrix Φ^* therefore equals the product of trait and method dispersions. With the Kelly & Fiske assessment data, both dispersion components are small, with 0.0122 and 0.0354, respectively, indicating that most of the systematic variance of the optimally-scaled ratings is due to the general variate δ_g. Method differences account for slightly more variation in the Assessment data than trait differences.

3.4.2 Canonical decomposition of covariance component blocks

Interpretation of covariance component estimates is greatly facilitated when the diagonal blocks $\Phi_{\tau\tau}^*$ and $\Phi_{\mu\mu}^*$ can be transformed into a diagonal structure. Choleski factorization and Eigenvalue decomposition are well-known traditional methods for this purpose.

Under Choleski factorization, a nonnegative definite symmetric matrix **A** is decomposed into the product of a lower triangular matrix **S** and its transpose: $\mathbf{A} = \mathbf{SS}'$, with $\mathbf{S}'\mathbf{S}$ diagonal. Computational procedures are described in many texts, for instance, Anderson (1984), Bock (1975), Finn (1974), and Maindonald (1984).

The results of Choleski factorization depend on the order of calculation. If **A** is of order $q \times q$, there will be $q!$ numerically different Choleski factors **S** with the equivalent product $\mathbf{SS}' = \mathbf{A}$. Order dependence of **S** may be put to an advantage when the contrast variates can be entered by importance or, in reverse order, by dubiosity. Then, s_{11}^2, the squared first diagonal entry in **S**, contains the relative variance due to the most important contrast, s_{22}^2 is the partial variance of the second most important contrast, adjusted for effects of the first one, s_{33}^2 the partial contribution of the third contrast, adjusted for the first two, and so on. These values may be evaluated in step-down fashion as successive partial contributions.

Unfortunately, there are many instances when an importance ranking of the latent variates is not meaningful on substantive grounds; the present analysis of the Kelly & Fiske assessment data being one of them. When the variates cannot be ranked beforehand, values in the diagonal of $\mathbf{S}$ are arbitrary and cannot be interpreted by themselves. Yet, even in this case, Choleski factorization will furnish a canonical matrix decomposition in the form $\mathbf{A} = \mathbf{S}\,\mathbf{I}\,\mathbf{S}'$, effectively reducing the covariance matrices $\Phi^*_{\tau\tau}$ and $\Phi^*_{\mu\mu}$ to orthonormal variates. The derived solution can then be further rotated to aid interpretation (see below).

Eigenvalue decomposition is another well-known method to describe a matrix in terms of a canonical structure. Eigenvalues λ_ℓ and the corresponding (non-zero) Eigenvectors $\mathbf{q}_\ell$ of a symmetric matrix $\mathbf{A}$ are defined as the roots of

$$\mathbf{A}\mathbf{q}_\ell = \mathbf{q}_\ell \lambda_\ell. \tag{36}$$

Solutions can be obtained by various numerical methods, many of which are implemented in such maintained software libraries as IMSL (IMSL, 1977), MATCAL (Bock & Repp, 1974), and the NAG library (NAG, Ltd., Oxford, U.K.).

Eigenvectors associated with different Eigenvalues of the same symmetric matrix are orthogonal. All Eigenvectors of $\mathbf{A}$ may be scaled to unit-length and assembled in the columns of the matrix $\mathbf{Q}$, so that $\mathbf{Q}'\mathbf{Q} = \mathbf{I}$. By collecting the associated Eigenvalues in the same order in the diagonal matrix $\mathbf{D}_\lambda$, Equation 36 can be written more compactly as

$$\mathbf{A}\mathbf{Q} = \mathbf{Q}\mathbf{D}_\lambda. \tag{37}$$

This further implies the canonical decomposition

$$\mathbf{A} = \mathbf{Q}\mathbf{D}_\lambda\mathbf{Q}'. \tag{38}$$

It has become customary to base the interpretation of Eigenanalysis on the weighted principal components $\mathbf{P} = \mathbf{Q}\mathbf{D}_\lambda^{1/2}$ rather than on the normalized components $\mathbf{Q}$.

The size of the Eigenvalues in $\mathbf{D}_\lambda$ reflects the variance of the respective Eigencomponents of $\mathbf{A}$: $\mathbf{q}_\ell'\mathbf{A}\mathbf{q}_\ell = \lambda_\ell$. It is well known that the largest Eigenvalue is the size of the largest variance component in $\mathbf{A}$, the second largest Eigenvalue the variance of the largest component that is orthogonal to the first, etc. (e.g., Anderson, 1984). The size of the Eigenvalue becomes a useful indicator for empirical importance of the principal components of a covariance matrix. Small variance components are likely redundant.

Since arbitrary choice of trait or method contrasts affects the estimate of Φ^*, computation of principal components must be based on the entire Gramian product $\mathbf{K}\Phi^*\mathbf{K}'$. In the block-diagonal CCA model, this covariance structure can be additively partitioned into general, trait, and method components as:

$$\mathbf{K}\Phi^*\mathbf{K}' = (\mathbf{K}_g|\mathbf{K}_\tau|\mathbf{K}_\mu)\left(\begin{array}{c|c|c} \sigma_g^{*2} & \multicolumn{2}{c}{\text{(symm.)}} \\ \hline 0 & \Phi^*_{\tau\tau} & \\ \hline 0 & 0 & \Phi^*_{\mu\mu} \end{array}\right)\left(\begin{array}{c} \mathbf{K}_g' \\ \hline \mathbf{K}_\tau' \\ \hline \mathbf{K}_\mu' \end{array}\right) \tag{39}$$

$$= \mathbf{K}_g\sigma_g^{*2}\mathbf{K}_g' + \mathbf{K}_\tau\Phi^*_{\tau\tau}\mathbf{K}_\tau' + \mathbf{K}_\mu\Phi^*_{\mu\mu}\mathbf{K}_\mu' \tag{40}$$

Table 10: Unrotated component loadings **P** of the Kelly & Fiske data

Method	Trait	δ_g	p_{t_1}	p_{t_2}	p_{t_3}	p_{t_4}	p_{m_1}	p_{m_2}
Staff Ratings	Assertive	.258	-.331	-.233	-.121	.020	-.079	-.103
	Cheerful	.258	-.312	.229	-.053	-.075	-.079	-.103
	Serious	.258	.733	-.006	-.063	-.018	-.079	-.103
	Poise	.258	-.045	.130	.051	.124	-.079	-.103
	Interests	.258	-.042	-.120	.190	-.047	-.079	-.103
Team-mate Ratings	Assertive	.258	-.331	-.233	-.121	.020	-.147	.084
	Cheerful	.258	-.312	.229	-.053	-.075	-.147	.084
	Serious	.258	.733	-.006	-.063	-.018	-.147	.084
	Poise	.258	-.045	.130	.051	.124	-.147	.084
	Interests	.258	-.042	-.120	.190	-.047	-.147	.084
Self Ratings	Assertive	.258	-.331	-.233	-.121	.020	.227	.018
	Cheerful	.258	-.312	.229	-.053	-.075	.227	.018
	Serious	.258	.733	-.006	-.063	-.018	.227	.018
	Poise	.258	-.045	.130	.051	.124	.227	.018
	Interests	.258	-.042	-.120	.190	-.047	.227	.018
Variance		1.0	2.246	.415	.182	.072	.239	.054

Separate Eigenstructures should be computed for the trait component $\mathbf{K}_\tau \Phi^*_{\tau\tau} \mathbf{K}'_\tau$ and the method component $\mathbf{K}_\mu \Phi^*_{\mu\mu} \mathbf{K}'_\mu$.

Table 10 shows the unrotated principal components of the Kelly & Fiske assessment data, computed from the block diagonal CCA solution. All seven variates are now uncorrelated and have unit variance, the columns of component loadings are weighted contrasts, sorted within blocks with respect to the explained variance. The first trait component, p_{t1} has more than twice the variance of the general variate. It is clearly defined as a contrast between *Seriousness* on one hand and the two variables *Assertiveness* and *Cheerfulness* on the other. *Seriousness* has the largest trait component—this variable is most distinct from the δ_g, the general variate, and, consequently, from most of the remaining traits in the study. *Assertiveness* and *Cheerfulness* are further removed from *Seriousness* than *Poise* and *Broad Interests*. The remaining trait components are relatively minor: p_{t2} through p_{t4} reflect some differences between the four trait domains other than *Seriousness*. Finally, the two method components indicate that most method variance is due to the difference between self ratings and ratings by others.

3.4.3 Blockwise VARIMAX rotation

Varimax rotated canonical components (Kaiser, 1958) comprise the same information as their unrotated counterparts and as the covariance component matrix

Table 11: Rotated component loadings **P** of the Kelly & Fiske data

Method	Trait	δ_g	p_{t_1}	p_{t_2}	p_{t_3}	p_{t_4}	p_{m_1}	p_{m_2}
Staff Ratings	Assertive	.258	-.407	-.078	-.030	.076	-.014	-.129
	Cheerful	.258	-.084	-.374	.064	-.086	-.014	-.129
	Serious	.258	.500	.481	.176	.171	-.014	-.129
	Poise	.258	.018	-.043	.019	-.185	-.014	-.129
	Interests	.258	-.025	.016	-.231	.022	-.014	-.129
Team-mate Ratings	Assertive	.258	-.407	-.078	-.030	.076	-.170	-.004
	Cheerful	.258	-.084	-.374	.064	-.086	-.170	-.004
	Serious	.258	.500	.481	.176	.171	-.170	-.004
	Poise	.258	.018	-.043	.019	-.185	-.170	-.004
	Interests	.258	-.025	.016	-.231	.022	-.170	-.004
Self Ratings	Assertive	.258	-.407	-.078	-.030	.076	.184	.134
	Cheerful	.258	-.084	-.374	.064	-.086	.184	.134
	Serious	.258	.500	.481	.176	.171	.184	.134
	Poise	.258	.018	-.043	.019	-.185	.184	.134
	Interests	.258	-.025	.016	-.231	.022	.184	.134

Φ^*, except that the rotation leads to a simple structure solution. Many students are familiar with simple structure solutions in factor analysis, from where it is a minor step to the interpretation of a simple structure derived from covariance components.

To demonstrate such a simple structure, the trait and method blocks in Table 10 were subjected to separate varimax rotations. The result is shown in Table 11. Evidence in the trait block now clearly shows that the four simple contrasts between *Seriousness* and each of the other traits are uncorrelated. The effects sizes of the rotated components corroborate earlier findings: *Assertiveness* and *Cheerfulness* are most disparate from *Seriousness*, while *Unshakable Poise* and *Broad Interests* are located somewhat closer. The rotated method components also appear as independent simple contrasts: The first component shows a difference between *self* ratings and *teammate* ratings, the second indicates that *self* ratings and *staff* ratings vary in different directions. Both rotated method component show about equal size.

3.5 Discussion

This paper studies the performance of two classes of multivariate linear model structures for the multitrait-multimethod matrix: *confirmatory factor analysis* and generalized *covariance component analysis*. Notable submodels with applications to multitrait-multimethod analysis are identified in each class. *Trait-only* and independent *trait-method decomposition* models are selected from factor analysis, *block-diagonal* and *independent-common-variation* represent the covariance component approach. The four models are partially nested and two strands of hierarchical model testing may be pursued. The two factor analytic models can be directly compared to each other and the statistical significance due to added method factors can be tested comparing the two model likelihoods. The other line of nested hierarchical model testing allows comparisons between the trait-only factor model and the two generalized CCA models.

Performance of all four model types is evaluated with 23 empirical MTMM matrices using the criteria of identification, convergence, admissibility, and model fit.

With these data, the trait-only factor model was generally found to be identified and converged, yet the solutions were often inadmissible, and model fit was typically very poor. However, this is a positive result compared to the trait-method factor model, which never even converged to an admissible solution. Formal evidence is provided showing that the trait-method factor is rotationally (and conceptually) underdetermined. This is bad news, because the trait-method decomposition model has been extensively promoted in the literature and its deficiencies are not yet widely known.

Two types of identification problems are found with generalized covariance component models. The first is trivial and easily removed: because the scale factors are estimated, the scale of the variance components is lost, making it necessary to fix a single variance component or one scaling constant at a non-zero value. Only the *relative* size of covariance component estimates is meaningful. The second identification problem is more severe: scale-free generalization of covariance component analysis finds its limitations when the common variate is allowed to correlate with the contrast components. In the presence of any kind of empirical sampling error; i.e., in all empirical applications, estimation of these parameter groups gives very unreliable results. Estimation of these correlations is only possible when the latent scale of the measured variables is assumed to be known (and vice versa). This new finding qualifies some of the very optimistic statements by Wiley et al. (1973).

Block-diagonal and independent-common-variation CCA converged to admissible solutions in about half the cases, with acceptable model fit for 5 or 6 of the 23 datasets. The fact that just a moderate number of MTMM matrices could be successfully modeled is a favorable result, considering that for several of these datasets multitrait-multimethod validation would have been questionable on substantive grounds already. A good statistical model must be falsifiable on empirical grounds to be of any practical use.

Despite its impressive estimation properties, covariance component analysis has not gained near as much popularity among researchers as the competing factor analytic model. One of the major reasons for this development may be that the interpretation of covariance components is difficult and unpopular. To facilitate interpretation, primary estimates of covariance components can be transformed into canonical variates and rotated into simple structure or into other orientations that may be heuristically helpful. Worked examples are provided.

Covariance component analysis provides a fundamental vehicle for the assessment of trait validity. Evaluation of validity should be based on a comprehensive model of the parameters that underlie an observed correlation structure, rather than the individual sample correlation components themselves. In reference to the treatment by Campbell & Fiske (1959) it is observed that *convergent validity* is reflected by disappearing method covariance structures; i.e., $\Phi^*_{\mu\mu} = 0$, while *discriminant validity* is established when the determinant of the trait covariance component matrix $\Phi^*_{\tau\tau}$ is large. The covariance structure approach has the interesting implication that traits can be validated in observational studies only insofar as they differ from other traits in the same study.

Appendix

Description of Datasets

Model performance is evaluated with 23 empirical correlation matrices published in the psychological and sociological literature. Nine matrices are taken from the original article by Campbell & Fiske (1959), the other 14 datasets were contributed in various papers written since. A synoptic characterization of each MTMM matrix is provided in Table 12.

The size of measurement design of the datasets varies from 2 traits by 2 methods to 4 traits by 5 methods and 15 traits by 3 methods. Sample sizes range between 35 and 750. Information on means and variances was notably absent from all reports, reflecting the traditional neglect of the scale of measurement in much of psychological research.

The trait domains are variously conceived as abilities, social dispositions or social behavior, attitudes or attitude components, drives, and social desirability judgments. Attractiveness of different methods as study objects has changed remarkably over the years: while Campbell & Fiske compared mostly effects of self ratings, ratings by others, and objective measures, later studies concern the effects of different question formats or of different panels engaging in political preference judgments.

Two studies show unorthodox method concepts. All six measures in the "Intelligence and Alertness (Thorndike)" data are paper-and-pencil assessments of ability with "Intelligence" and "Mental Alertness" labeled as traits and "Memory", "Comprehension", and "Vocabulary" labeled as methods. The "Clinical Clerkships (Boodoo)" data assess general dispositions in "Pediatrics", "Internal Medicine", and "Surgery" as traits, based on the performance in such method domains as "Skills", "Problem solving", etc.. In these two cases, both facets of the measurement design are of substantive interest and the "trait" versus "method" distinction becomes arbitrary. Boodoo (1985) did, in fact, label her facets the exact opposite way. One should keep in mind that multivariate methods just as easily accommodate cross-classifications among several trait facets as they can handle measurement designs of traits by methods. Hierarchical model testing is aided from a substantive point of view when the facets clearly differ in relevance, but estimation methods and fit statistics remains unaffected even when all facets are equally important.

One of the correlation matrices, the "Moral Dilemma (Shepherd)" data, proved not to be positive definite and, consequently, not Gramian. This may have been due to a typesetting error or to some pairwise deletion of missing data remaining unreported. For the present analyses, the originally published matrix was smoothed subtracting the negative roots from the correlation structure and adding a ridge of small variance components to the diagonal.

Table 12: Description of datasets

Name, Trait-Conception	Size			Domains		Data Source
	Traits	Methods	N	Traits	Methods	
Intelligence and Effort (Mayo) **Abilities**	2	2	166	1 Intelligence 2 Effort	a Peer Ratings b Objective Measures	Campbell and Fiske (1959)
Intelligence and Alertness (Thorndike) **Abilities**	2	3	750	1 Intelligence 2 Mental Alertness	a Memory b Comprehension c Vocabulary	Campbell and Fiske (1959)
Popularity and Expansiveness (Borgatta) **Social Dispositions**	2	4	125	1 Popularity 2 Expansiveness	a Sociometr. Self Rating b Rating by Others c Observed Group Interaction d Observed Role Playing	Campbell and Fiske (1959)
Smoking and Capital Punishment (Jaccard, et al.) **Attitudes**	2	4	35	1 Cigarette Smoking 2 Capital Punishment	a Semantic Differential b Likert Scaling c Thurstone Scaling d Guilford Scaling	Jaccard, Weber, and Lundmark (1975)
Leadership (Summers, et al.) **Social Dispositions**	2	4	290	1 Community Leadership 2 Educational Leadership	Panels of a 17 School Leaders b 20 Organization Heads c 19 Popular Judges d 196 Heads of Households	Summers, Seiler, and Wiley (1970)

Table 12—*Continued*

Name, Trait-Conception	Size: Traits	Size: Methods	Size: N		Domains: Traits		Domains: Methods	Data Source
Attitudes	3	2	57		Attitudes			Campbell and
to Authority				1	To father	a	Interview	Fiske (1959)
(Burwen and				2	To boss	b	Check	
Campbell)				3	To peer		List	
Attitudes								
Drives in	3	2	50	1	Hunger	a	Obstruc-	Campbell and
Rats				2	Thirst		tion Box	Fiske (1959)
(Anderson)				3	Sex	b	Activity	
Drives								
Involvement	3	3	96		Involvement		Rating Scale	Arora (1982)
Components				1	Situa-	a	Stapel	
(Arora)					tional	b	Likert	
Attitudes				2	Enduring	c	Semantic	
				3	Response		Differen-	
							tial	
Job	3	3	149	1	Getting	a	Peer	Dickinson &
Behavior					along with		Nominations	Tice (1973)
(Dickinson					others	b	Peer Check-	
& Tice)				2	Dedication		list	
Social				3	Ability to		Ratings	
Disposi-					apply	c	Self Check-	
tions					learning		list	
							Ratings	
Three	3	3	105		Attitude	a	Likert	Flamer (1983)
Attitudes					Towards		Scales	
(Flamer,				1	Disci-	b	Thurstone	
Sample 1)					pline of		Scales	
Attitudes					children	c	Semantic	
				2	Mathematics		Differ-	
				3	The law		ential	
Three	3	3	105		Attitude	a	Likert	Flamer (1983)
Attitudes					Towards		Scales	
(Flamer,				1	Disci-	b	Thurstone	
Sample 2)					pline of		Scales	
Attitudes					children	c	Semantic	
				2	Mathematics		Differ-	
				3	The law		ential	

Table 12—*Continued*

Name, Trait-Conception	Size			Domains		Data Source
	Traits	Methods	N	Traits	Methods	
Stress Measures (Karst and Most) **Arousal States**	3	3	80	Stress 1 Anticipatory 2 During Performance 3 Post Performance	a General Self Ratings b Anchored Self Ratings c Finger Sweat Print	Karst and Most (1973)
Job Performance (Lawler) **Social Dispositions**	3	3	113	1 Quality of job performance 2 Ability to perform job 3 Effort put forth on the job	Ratings by a Superiors b Peers c Self	Lawler (1967)
Moral Dilemma (Shepherd) **Beliefs and Attitudes**	3	3	487	Morality 1 Negative 3 Positive 3 Achievement of	Three different test forms	Shepherd (1977)
Attitudes to Contraceptives (Kothandapani) **Attitude Components**	3	4	100	Attitude Components 1 Affective 2 Behavioral 3 Cognitive	a Thurstone Scaling b Likert Scaling c Guttman Scaling d Guilford Scaling	Kothandapani (1971)

Table 12—*Continued*

Name, Trait-Conception	Size			Domains		Data Source
	Traits	Methods	N	Traits	Methods	
Attitudes to the Church (Ostrom) **Attitude Components**	3	4	189	Attitude Components 1 Affective 2 Behavioral 3 Cognitive	a Thurstone Scaling b Likert Scaling c Guttman Scaling d Guilford Scaling	Ostrom (1969)
Drug Use Reports (Stacy et al.) **Social Behavior**	3	4	190	1 Alcohol 2 Marijuana 3 Nicotine	a Self Rating b Self Intake Report c Peer Rating d Peer Intake Report	Stacy et al. (1985)
Clinical Clerkships (Boodoo) **Social Dispositions**	3	5	136	1 Pediatrics 2 Internal Medicine 3 Surgery	Teacher Ratings a Skills b Problem Solving c Relationships d Knowledge e Attitude	Boodoo (1985)
Personality Traits (Kelly and Krey) **Social Dispositions**	4	2	311	Social Traits 1 Courtesy 2 Honesty 3 Poise 4 School Drives	a Peer Rating b Association Test	Campbell and Fiske (1959)

Table 12—*Continued*

Name, Trait-Conception	Size: Traits	Size: Methods	N	Domains: Traits	Domains: Methods	Data Source
Desirability and Frequency Ratings of Personality Traits (Jackson and Singer) **Judgments on Personality Traits**	4	5	480	Judgments on Personality Traits 1 Femininity 2 Anxiety 3 Somatic Complaints 4 Socially Deviant Attitudes	a Desirable in Self b Desirable in Others c What Others Find Desirable d Frequency of Occurrence e Harmfulness	Jackson (1975)
Interaction Process Variables (Borgatta) **Social Behavior**	5	3	125	Social Behavior 1 Shows Solidarity 2 Gives Suggestion 3 Gives Opinion 4 Gives Orientation 5 Shows Disagreement	a Free Behavior b Role Playing c Projective Test	Campbell and Fiske (1959)
Guilford-Martin Factors (Carroll) **Personality Dispositions**	5	3	110	1 S 2 T 3 D 4 C 5 -R	a Inventory b Self Rating c Peer Rating	Campbell and Fiske (1959)
Clinical Assessment (Kelly and Fiske) **Social Traits**	5	3	124	1 Assertive 2 Cheerful 3 Serious 4 Unshakable Poise 5 Broad Interests	a Staff Rating b Teammate Rating c Self Rating	Campbell and Fiske (1959)

References

Althauser, R.P. (1974) Inferring validity from the multitrait-multimethod matrix: another assessment. In H.L. Costner [Ed.] *Sociological Methodology 1973–1974*. San Francisco: Jossey-Bass

Althauser, R.P. & Heberlein, T.A. (1970) Validity and the multitrait-multimethod matrix. In E.F. Borgatta [Ed.] *Sociological Methodology 1970*. San Francisco: Jossey-Bass

Althauser, R.P., Heberlein, T.A., & Scott, R.A. (1971) A causal assessment of validity: the augmented multitrait-multimethod matrix. In H.M. Blalock [Ed.] *Causal models in the social sciences*. Chicago: Aldine

Anderson, T.W. (1984) *An introduction to multivariate statistical analysis* 2nd. Ed.). New York: Wiley

Anderson, T.W. & Rubin, H. (1956) Statistical inference in factor analysis. *Proceedings of the Third Berkeley Symposium of Mathematical Statistics and Probability*, **5**,111–150

Arora, R. (1982) Validati. of an S-O-R model for situation, enduring, and response components of involvement. *Journal of Marketing Research*, **19**, 505–516

Beals, R., Krantz, D.H., & Tversky, A. (1968) Foundations of multidimensional scaling. *Psychological Review*, **75**, 127–142

Bock, R.D. (1960) Components of variance analysis as a structural and discriminal analysis of psychological tests. *The British Journal of Statistical Psychology*, **13**(2), 151–163

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

Bock, R.D. & Bargmann, R.E. (1966) Analysis of covariance structures. *Psychometrika*, **31**(4), 507–534

Bock, R.D., Dicken, C., & Van Pelt, J. (1969) Methodological implications of content-acquiescence correlation in the MMPI. *Psychological Bulletin*, **71**(2), 127–139

Boodoo, G.M. (1985) The structure of clinical evaluation forms across three clerkships. *Paper presented at the American Educational Research Association Meeting, Chicago, April 1985*.

Browne, M.W. (1984a) Asymptotically distribution-free methods for the analysis of covariance structures. *British Journal of Mathematical and Statistical Psychology*, **37**, 1–21

Browne, M.W. (1984b) The decomposition of multitrait-multimethod matrices. *British Journal of Mathematical and Statistical Psychology*, **37**, 62–83

Campbell, D.T. (1969) Definitional multiple operationalism. *et al.*, **2**, 14–17

Campbell, D.T. & Fiske, D.W. (1959) Convergent and discriminant validation by the multitrait-multimethod matrix. *Psychological Bulletin*, **56**, 81–105

Dickinson, T.L. & Tice, T.E. (1973) A multitrait-multimethod analysis of scales developed by retranslation. *Organizational Behavior and Human Perfor-*

mance, **9**, 421–438

Finn, J.E. (1973) A note on a sufficiency condition for uniqueness of a restricted factor matrix. *Psychometrika*, **38**, 141–143

Finn, J.D. (1974) *A general model for multivariate data analysis*. New York: Holt, Rinehart & Winston

Fiske, D.T. (1982) Convergent-discriminant validation in measurement and research strategies. *New Directions for Methodology of Social and Behavioral Science*, **12**, 77–92

Flamer, S. (1983) Assessment of the multitrait-multimethod matrix validity of Likert scales via confirmatory factor analysis. *Multivariate Behavioral Research*, **18**, 275–308

Graybill, F.A. (1961) *An introduction to Linear Statistical Models*. Vol. 1. New York: McGraw-Hill

Jaccard, J., Weber, J., & Lundmark, J. (1975) A multitrait-multimethod analysis of four assessment procedures. *Journal of Experimental Social Psychology*, **11**, 149–154

Jackson, D.N. (1975) Multimethod factor analysis: a reformulation. *Multivariate Behavioral Research*, **10**, 259–275

Jennrich, R.I (1978) Rotational equivalence of factor loading matrices with specified values. *Psychometrika*, **43**, 421–426

Jöreskog, K.G. (1966) Testing a simple structure hypothesis in factor analysis. *Psychometrika*, **31**, 165–178

Jöreskog, K.G. (1971) Statistical analysis of sets of congeneric tests. *Psychometrika*, **36**(2), 109–133

Jöreskog, K.G. (1978) Structural analysis of covariance and correlation matrices. *Psychometrika*, **43**(4), 443–477

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

Jöreskog, K.G. & Sörbom, D. (1986) *LISREL 6. Analysis of Linear Structural Relationships by the Method of Maximum Likelihood. User's guide.* Mooresville, IN: Scientific Software

Kalleberg, A.L. & Kluegel, J.R. (1975) Analysis of the multitrait-multimethod matrix: some limitations and an alternative. *Journal of Applied Psychology*, **60**(1), 1–9

Karst, T.O. & Most, R. (1973) A comparison of stress measures in an experimental analogue to public speaking. *Journal of Consulting and Clinical Psychology*, **4**(3), 342–348

Kothandapandi, V. (1971) Validation of feeling, belief, and intention to act as three components of attitude and their contribution to prediction of contraceptive behavior. *Journal of Personality and Social Psychology*, **19**(3), 321–333

Lawler, E.E., III (1967) The multitrait-multimethod approach to measuring managerial job performance. *Journal of Applied Psychology*, **51**(5), 369–381

Lawley, D.N. (1940) The estimation of factor loadings by the method of maximum likelihood. *Proceedings of the Royal Statistical Society of Edinburgh*, **60**, 64–82

Lawley, D.N. & Maxwell, A.E. (1971) *Factor analysis as a statistical method*. 2nd ed. London: Butterworth

Lord, F.M. & Novick, M.R. (1968) *Statistical theories of mental test scores*. Reading, MA: Addison-Wesley

Maindonald, J.H. (1984) *Statistical Computation*. New York: Wiley

Ostrom, T.M. (1969) The relationship between affective, behavioral, and cognitive components of attitude. *Journal of Experimental Social Psychology*, **5**, 12–30

Schmitt, N. (1978) Path analysis of multitrait-multimethod matrices. *Applied Psychological Measurement*, **2**(2), 157–173

Searle, S.R. (1982) *Matrix algebra useful for statistics*. New York: Wiley

Shepherd, G. (1977) A critique and extension of Bronfenbrenner's moral dilemma test. *International Journal of Psychology*, **12**(3), 207–217

Stacy, A.W., Widaman, K.F., Hays, R., & DiMatteo, M.R. (1985) Validity of self-reports of alcohol and other drug use: A multitrait-multimethod assessment. *Journal of Personality and Social Psychology*, **49**(1), 219–232

Summers, G.F., Seiler, L.H. & Wiley, G. (1970) Community and educational leadership. In E.F. Borgatta [Ed.], *Sociological Methodology*. San Francisco: Jossey-Bass

Werts, C.E., Jöreskog, K.G. & Linn, R.L. (1972) A multitrait-multimethod model for studying growth. *Educational and Psychological Measurement*, **32**, 655–678

Werts, C.E. & Linn, R.L. (1970) Path analysis. Psychological examples. *Psychological Bulletin*, **74**(3), 193–212

Werts, C.E., Linn, R.L., & Jöreskog, K.G. (1971) Estimating the parameters of path models involving unmeasured variables. In H.M. Blalock [Ed.] *Causal models in the social sciences*. Chicago: Aldine

Wiley, D., Schmidt, W.H., & Bramble, W.J. (1973) Studies of a class of covariance structure models. *Journal of the American Statistical Association*, **68**(342), 317–323

Wothke. W. (1984) *The estimation of trait and method components in multitrait-multimethod measurement*. University of Chicago, Department of Behavioral Science: Unpublished Dissertation.