DOCUMENT RESUME

ED 053 157 TM 000 666

AUTHOR	Werts, Charles E.; And Others
TITLE	A Multitrait-Multimethod Model for Studying Growth.
INSTITUTION	Educational Testing Service, Princeton, N.J.
SPONS AGENCY	Department of Health, Education, and Welfare, Washington, D.C.; Office of Education (DHEW), Washington, D.C.
REPORT NO	RB-71-13
PUB DATE	Mar 71
GRANT	OEG-2-700033(509)
NOTE	36p.
EDRS PRICE	EDRS Price MF-$0.65 HC-$3.29
DESCRIPTORS	Analysis of Covariance, Analysis of Variance, Correlation, *Factor Analysis, *Factor Structure, Hypothesis Testing, *Mathematical Models, *Mathematics, *Personal Growth, Statistics

ABSTRACT

The logical structure of the Campbell and Fiske multitrait-multimethod approach is applied to the problem of studying growth and its determinants. The resulting model is a special case of Joreskog's general model for the analysis of covariance structures. The relationships of traditional psychometric formulations to this model are detailed. (Author)

RB-71-13

A MULTITRAIT-MULTIMETHOD MODEL FOR STUDYING GROWTH

Charles E. Werts, Karl G. Jöreskog
and Robert L. Linn

This Bulletin is a draft for interoffice circulation. Corrections and suggestions for revision are solicited. The Bulletin should not be cited as a reference without the specific permission of the authors. It is automatically superseded upon formal publication of the material.

Educational Testing Service
Princeton, New Jersey
March 1971

A MULTITRAIT-MULTIMETHOD MODEL FOR STUDYING GROWTH

Abstract

The logical structure of the Campbell and Fiske (1959) multitrait-multimethod approach is applied to the problem of studying growth and its determinants. The resulting model is a special case of Jöreskog's (1970a) general model for the analysis of covariance structures. The relationships of traditional psychometric formulations to this model are detailed.

A MULTITRAIT-MULTIMETHOD MODEL FOR STUDYING GROWTH[1]

Charles E. Werts, Karl G. Jöreskog, and Robert L. Linn

Werts and Linn (1970a) have suggested that a multitrait-multimethod approach (Campbell & Fiske, 1959) might be used for studying growth. The purpose of this paper is to detail such a model and to outline implications for the study of growth. The major focus of our exposition will be the logic of this model rather than the estimation of parameters or testing the fit of the model to data. A comprehensive discussion of appropriate estimation and fit-testing procedures may be found in Jöreskog (1970a), whose general model for the analysis of covariance structures subsumes the models used in this paper.

The Model

The multitrait-multimethod approach may be treated as a problem in confirmatory factor analysis (Jöreskog, 1970a, 1971). For illustrative purposes we will consider the example of three traits and three methods since this is the minimum number of traits and methods required to produce unique (defined in Jöreskog, 1969, pp. 185-186) parameter estimates, given the assumption that each observed measure loads on only one trait and one method factor and all factors are oblique. The general factor analytic model is:

$$\underset{\sim}{y} = \underset{\sim}{\mu} + \underset{\sim\sim}{\Lambda T} + \underset{\sim}{e} \tag{1}$$

[1] The research reported herein was performed pursuant to Grant No. OEG-2-700033(509) with the United States Department of Health, Education, and Welfare and the Office of Education.

where $\underset{\sim}{y}$ is the vector of observed scores,

$\underset{\sim}{\mu}$ is the mean vector of $\underset{\sim}{y}$,

$\underset{\sim}{\Lambda}$ is a matrix of factor loadings,

$\underset{\sim}{T}$ is a vector of common factor scores, and

$\underset{\sim}{e}$ is a vector of unique factor scores corresponding to specific factors and/or errors of measurement.

For our example:

$$\underset{\sim}{y}' = (y_{11}, y_{21}, y_{31}, y_{12}, y_{22}, y_{32}, y_{13}, y_{23}, y_{33}) \quad , \tag{1a}$$

where in y_{ij}, i = method and j = trait,

$$\underset{\sim}{T}' = (T_1, T_2, T_3, M_1, M_2, M_3) \quad , \tag{1b}$$

where T_j = the j -th trait factor,

M_i = the i -th method factor,

$$\underset{\sim}{\Lambda} = \begin{bmatrix} A_{11} & 0 & 0 & B_{11} & 0 & 0 \\ A_{21} & 0 & 0 & 0 & B_{21} & 0 \\ A_{31} & 0 & 0 & 0 & 0 & B_{31} \\ 0 & A_{12} & 0 & B_{12} & 0 & 0 \\ 0 & A_{22} & 0 & 0 & B_{22} & 0 \\ 0 & A_{32} & 0 & 0 & 0 & B_{32} \\ 0 & 0 & A_{13} & B_{13} & 0 & 0 \\ 0 & 0 & A_{23} & 0 & B_{23} & 0 \\ 0 & 0 & A_{33} & 0 & 0 & B_{33} \end{bmatrix} \tag{1c}$$

where A_{ij} are loadings on trait factors and

B_{ij} are loadings on method factors.

The expected variance-covariance matrix $\underset{\sim}{\Sigma}$ of $\underset{\sim}{y}$ is then given by

$$\underset{\sim}{\Sigma} = \underset{\sim}{\Lambda}\underset{\sim}{\Phi}\underset{\sim}{\Lambda}' + \theta^2 \tag{2}$$

where θ^2 is a diagonal matrix whose elements are the variances of $\underset{\sim}{e}$.
Since all factors are oblique, in our example:

$$\Phi = \begin{bmatrix} V_{T_1} & & \text{Symmetric} & & & \\ C_{T_1T_2} & V_{T_2} & & & & \\ C_{T_1T_3} & C_{T_2T_3} & V_{T_3} & & & \\ C_{T_1M_1} & C_{T_2M_1} & C_{T_3M_1} & V_{M_1} & & \\ C_{T_1M_2} & C_{T_2M_2} & C_{T_3M_2} & C_{M_1M_2} & V_{M_2} & \\ C_{T_1M_3} & C_{T_2M_3} & C_{T_3M_3} & C_{M_1M_3} & C_{M_2M_3} & V_{M_3} \end{bmatrix} \tag{2a}$$

where the C 's are covariances and the V 's are variances.

Following Jöreskog (1970a), parameters will be labelled as one of three kinds: (1) fixed parameters that have been assigned given values; (2) constrained parameters that are unknown but equal to one or more other parameters; and (3) free parameters that are unknown and not constrained to be equal to any other parameter. The term "identifiable" will be used in the sense defined by Fisher (1966, p. 25): "we shall speak of that equation as identifiable (or identified) if there exists some combination of prior and posterior information which will enable us to distinguish its parameters from those of any other equation in the same form." For the models studied in this paper, the term "identifiable" is synonymous with the factor analyst's term "unique solution," i.e., a solution is "unique" if all linear transformations of the factors that leave the fixed parameters unchanged also leave the free parameters unchanged. As Jöreskog (1970b) notes: "Before an attempt is made to estimate a model of this kind, the identification problem must be examined." The number of overidentifying restrictions on the model is frequently of interest, for example, after standardizing factor variances (i.e., $V_{T_j} = V_{M_i} = 1$) the three method by three trait model has three overidentifying restrictions, i.e., $\underset{\sim}{\Sigma}$ has 45 distinct variances and covariances as compared to 42 free parameters to be estimated (18 factor loadings, 15 factor covariances in Φ, and nine residual variances in θ). The number of overidentifying restrictions is the degrees of freedom (df) for the test statistic in Jöreskog's general model (1970a, p. 241, sec. 1.4). The "path analysis" approach used by Werts and Linn (1970a) can be very useful in exploring the identification question in overidentified models. However, as noted by Hauser and Goldberger (1970) the "path analysis" literature does not adequately deal with the estimation problem in overidentified models, in part because the sample-population distinction is blurred.

The multitrait-multimethod approach considered above does not consider any functional relationships among the trait factors, i.e., the approach deals only with errors of measurement. In the study of growth, these trait factors correspond to initial status, final status, and the determinants of growth and a structural model showing the relationship among these variables must be specified. Substantive inferences about growth are based on estimates of the parameters of the structural model.

Suppose that the structural model for growth took the form:

$$T_3 = D_1 T_1 + D_2 T_2 + \xi \qquad (3)$$

where T_3 is the final status, T_2 is the initial status, and T_1 is a determinant of growth; all other influences on growth (represented by ξ) being independent of T_1 and T_2. In this model the initial status T_2 may influence the rate of growth. The parameters of equation (3) are just identifiable in terms of the elements of Φ, i.e., the number of restrictions on the overall model is not changed. Assuming that T_3 and T_2 are measurements on the same dimension as implied by the terms "initial" and "final" status, growth (Δ) is equal to $T_3 - T_2$. Werts and Linn (1970b) have shown that the regression weights for T_1 and T_2 are:

$$D_1 = D_{\Delta T_1 \cdot T_2} \quad , \qquad (4)$$

and

$$D_2 = 1 + D_{\Delta T_2 \cdot T_1} \qquad (5)$$

where $D_{\Delta T_1 \cdot T_2}$ is the regression weight of Δ on T_1 with T_2 controlled and $D_{\Delta T_2 \cdot T_1}$ is the regression weight of Δ on T_2 with T_1 controlled. In other words D_1 represents the direct influence of T_1 on growth and D_2 represents the direct influence of initial status on growth plus unity (which represents that part of T_3 which is initial status). Since $T_3 = \Delta + T_2$, substituting equations (4) and (5) into (3) yields:

$$\Delta = D_{\Delta T_1 \cdot T_2} T_1 + D_{\Delta T_2 \cdot T_1} T_2 + \xi \quad . \qquad (6)$$

In terms of T_1 , T_2 and ξ , equations (1b), (1c), and (2a) become:

$$\underset{\sim}{T}^* = (T_1, T_2, \xi, M_1, M_2, M_3) \quad , \tag{7a}$$

$$\underset{\sim}{\Lambda}^* = \begin{bmatrix} A_{11} & 0 & 0 & B_{11} & 0 & 0 \\ A_{21} & 0 & 0 & 0 & B_{21} & 0 \\ A_{31} & 0 & 0 & 0 & 0 & B_{31} \\ 0 & A_{12} & 0 & B_{12} & 0 & 0 \\ 0 & A_{22} & 0 & 0 & B_{22} & 0 \\ 0 & A_{32} & 0 & 0 & 0 & B_{32} \\ A_{13}D_1 & A_{13}D_2 & A_{13} & B_{13} & 0 & 0 \\ A_{23}D_1 & A_{23}D_2 & A_{23} & 0 & B_{23} & 0 \\ A_{33}D_1 & A_{33}D_2 & A_{33} & 0 & 0 & B_{33} \end{bmatrix} \quad , \tag{7b}$$

and

$$\Phi^* = \begin{bmatrix} 1 & & & \text{Symmetric} & & \\ C_{T_1T_2} & 1 & & & & \\ 0 & 0 & V_\xi & & & \\ C_{T_1M_1} & C_{T_2M_1} & C_{\xi M_1} & 1 & & \\ C_{T_1M_2} & C_{T_2M_2} & C_{\xi M_2} & C_{M_1M_2} & 1 & \\ C_{T_1M_3} & C_{T_2M_3} & C_{\xi M_3} & C_{M_1M_3} & C_{M_2M_3} & 1 \end{bmatrix} \quad , \tag{7c}$$

respectively. If the analyst wished to scale a factor by the unit of a particular measure this may be accomplished by setting the A_{ij} slope for the measure equal to unity (in which case the variance of the corresponding factor should not be standardized but left free to be estimated by the program). The assumption that T_2 and T_3 are measures on the same dimension is equivalent to setting the same method regression weights equal, i.e., in our example $A_{12} = A_{13}$, $A_{22} = A_{23}$, and $A_{32} = A_{33}$. As detailed by Werts and Linn (1970a) the effect of these restrictions is that the ratio of the variance of T_3 to T_2 is fixed. For estimation purposes it is convenient to standardize all factors except T_3 whose variance is fixed in relation to T_2. The model defined by equations (7a), (7b), and (7c) is no longer a simple factor analysis model, but may be estimated using Jöreskog's (1970a) general model for the analysis of covariance structures. For this purpose $\underset{\sim}{\Lambda}^*$ may be rewritten as the product of two matrices:

$$\underset{\sim}{\Lambda}^* = \underset{\sim}{B}\underset{\sim}{\Lambda}^{**}$$

where

$$\underset{\sim}{B} = \begin{bmatrix} 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & A_{13} & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & A_{23} & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & A_{33} \end{bmatrix}$$

and

$$\Lambda^{**} = \begin{bmatrix} A_{11} & 0 & 0 & B_{11} & 0 & 0 \\ A_{21} & 0 & 0 & 0 & B_{21} & 0 \\ A_{31} & 0 & 0 & 0 & 0 & B_{31} \\ 0 & A_{12} & 0 & B_{12} & 0 & 0 \\ 0 & A_{22} & 0 & 0 & B_{22} & 0 \\ 0 & A_{32} & 0 & 0 & 0 & B_{32} \\ D_1 & D_2 & 1 & X_{13} & 0 & 0 \\ D_1 & D_2 & 1 & 0 & X_{23} & 0 \\ D_1 & D_2 & 1 & 0 & 0 & X_{33} \end{bmatrix}$$

and $X_{13} = B_{13}/A_{13}$, $X_{23} = B_{23}/A_{23}$, $X_{33} = B_{33}/A_{33}$. By substitution:

$$\underset{\sim}{\Sigma} = \underset{\sim}{B}\underset{\sim}{\Lambda}^{**}\underset{\sim}{\Phi}^{*}\underset{\sim}{\Lambda}^{**'}\underset{\sim}{B}' + \underset{\sim}{\theta}^2 \quad ,$$

which is a special case of Jöreskog's (1970a) general model.

In using the computer program (Jöreskog, Gruvaeus, & van Thillo, 1970) the parameters A_{12} , A_{22} , A_{32} in Λ^{**} should be constrained to be equal to A_{13} , A_{23} , and A_{33} respectively in $\underset{\sim}{B}$. The resulting model has 45 distinct variances and covariances in $\underset{\sim}{\Sigma}$ and 40 free and constrained parameters (17 in $\underset{\sim}{\Lambda}^{**}$, 14 in $\underset{\sim}{\Phi}^{*}$, 9 in $\underset{\sim}{\theta}$, none in $\underset{\sim}{B}$ because of equality restraints), which means that the model has five overidentifying restrictions (df). The advantage of casting the analysis in terms of Jöreskog's general model is that, given the assumption that the observed variables are distributed normally, various hypotheses about the model may be tested in large samples.

In particular, we may wonder if trait factors are uncorrelated with methods factors and methods factors with each other as assumed by Cronbach and Furby (1970) and Werts and Linn (1970a) in their analysis of growth. To make this test, the analysis would be run with the model of (1a), (1b), and (1c), and (2a) with $V_{T_1} = V_{T_2} = V_{T_3} = V_{M_1} = V_{M_2} = V_{M_3} = 1$ and then the analysis would be made with $C_{T_1M_1} = C_{T_1M_2} = C_{T_1M_3} = C_{T_2M_1} = C_{T_2M_2} = C_{T_2M_3} = C_{T_3M_1} = C_{T_3M_2} = C_{T_3M_3} = C_{M_1M_2} = C_{M_1M_3} = C_{M_2M_3} = 0$. For our example, the initial analysis would yield a chi-square with three df for testing the fit of the model to the data. The second analysis would yield a chi-square with 15 df since 12 additional restrictions have been made. The increase in chi-square with 12 df is a test of the tenability of the additional restrictions. Starting with the same initial model, the tenability of assuming that $A_{12} = A_{13}$, $A_{22} = A_{23}$, and $A_{32} = A_{33}$ may be tested (dropping the $V_{T_3} = 1$ assumption) using the increase in chi-square with 2 df. Likewise starting with these assumptions (i.e., equations (7a), (7b), and (7c), and df = 5) hypotheses about growth can be tested, e.g., D_1' can be set equal to zero and the resulting change in χ^2 (df = 1) is a test of whether T_2 directly influences growth. To test whether initial status directly influences growth (i.e., whether $D_{\Delta T_2 \cdot T_1} = 0$), D_2 would be set equal to unity (see equation (4)), the increase in χ^2 (df = 1) testing this hypothesis. The fit of the observed variance-covariance matrix $\underset{\sim}{S}$ to the estimated elements of $\underset{\sim}{\Sigma}$ may be used to form some judgment as to changes in fit resulting from additional restrictions, especially when the χ^2 test is inappropriate because the assumption of multivariate normality is not reasonable.

As originally conceived by Campbell and Fiske (1959) the multitrait-multimethod approach required each trait to be measured with each method, as in the example analyzed above. The linear structural model approach proposed herein requires that model parameters be identifiable, a question which is unrelated to whether each trait is measured with each method. In order to fix the ratio of the variance of the final status to the initial status factor, only one pair of initial and final measures with the same units of measurement are required, i.e., the three sets of initial-final measures in our example serve to overidentify this variance ratio. The identification problem would be greatly simplified if one of these same method sets were replaced with different method measures, even though the resulting matrix would no longer be in the form required by Campbell and Fiske. Campbell and Fiske's argument that different method measures of a trait are required to improve convergent validity appears fundamentally sound and is a basic premise in our analysis. We have abandoned the particular type of analysis used by Campbell and Fiske because it fails to specify the underlying structure being postulated, and does not allow for nonsymmetrical method-by-trait combinations.

Relationship to Classical Test Theory

The multitrait-multimethod formulation can be shown to include various procedures derived from classical test theory as special cases, e.g., the commonly used formulas for reliability of differences, correlation of true initial status with true gain, and the correlation of true scores over time can be derived from the multitrait-multimethod model by imposing specifiable restrictions. To illustrate this point we shall examine the case of two parallel measures (y_{12}, y_{22}) given initially and two finally (y_{13}, y_{23}). First let us consider the analysis given the traditional assumptions that all errors of measurement are independent of each other and of the true scores. In our formulation this is equivalent to asserting that there are no methods factors. Without further assumptions the model may be represented in terms of equation (1) as

$$\underset{\sim}{y} = (y_{12}, y_{22}, y_{13}, y_{23}) \quad , \tag{8a}$$

$$\underset{\sim}{T} = (T_2, T_3) \quad , \tag{8b}$$

$$\underset{\sim}{\Lambda} = \begin{bmatrix} A_{12} & 0 \\ A_{22} & 0 \\ 0 & A_{13} \\ 0 & A_{23} \end{bmatrix} \quad , \tag{8c}$$

$$\underset{\sim}{\Phi} = \begin{bmatrix} V_{T_2} & \\ C_{T_2T_3} & V_{T_3} \end{bmatrix} \quad , \tag{8d}$$

and

$$\underset{\sim}{\theta}^2 = \begin{bmatrix} V_{e_{12}} & & & \\ 0 & V_{e_{22}} & & \\ 0 & 0 & V_{e_{13}} & \\ 0 & 0 & 0 & V_{e_{23}} \end{bmatrix} \quad . \tag{8e}$$

Assuming that initial and final status are on the same scale, "parallel" test assumptions are equivalent to (Jöreskog, 1971) fixing $A_{12} = A_{22} = A_{13} = A_{23} = 1$ and constraining $V_{e_{12}} = V_{e_{22}}$ and $V_{e_{13}} = V_{e_{23}}$. All parameters are identifiable and df = 5. Identification still occurs without the error variance assumptions (df = 3), i.e., in true score lexicon, "essentially tau-equivalent" measures (Lord & Novick, 1968, pp. 47-50) would suffice. If we choose to use nonparallel or "congeneric" (Jöreskog, 1971) measures, one pair of measures over time being on the same scale (e.g., $A_{12} = A_{13}$), V_{T_2} could be arbitrarily standardized (= 1), yielding an identifiable model with df = 1. In all these cases, growth statistics may be obtained from the parameter estimates or the model can be transformed to obtain growth statistics directly. Inserting $T_3 = T_2 + \Delta$ then:

$$T^* = (T_2 \ , \ \Delta) \quad , \tag{9a}$$

$$\underset{\sim}{\Lambda}^* = \begin{bmatrix} A_{12} & 0 \\ A_{22} & 0 \\ A_{13} & A_{13} \\ A_{23} & A_{23} \end{bmatrix} \quad , \tag{9b}$$

where $A_{12} = A_{13}$ by assumption, and

$$\underset{\sim}{\Phi}^* = \begin{bmatrix} V_{T_2} & \\ C_{T_2\Delta} & V_{\Delta} \end{bmatrix} \quad , \tag{9c}$$

where $V_{T_2} = 1$ for convenience.

Relevant growth statistics are:

$$\hat{\rho}_{T_2\Delta} = \text{correlation of initial status with gain} = \hat{C}_{T_2\Delta} \div \sqrt{\hat{V}_\Delta \hat{V}_{T_2}} \quad , \tag{10a}$$

$$\hat{D}_{\Delta T_2} = \hat{C}_{T_2\Delta} \div \hat{V}_{T_2} \quad , \tag{10b}$$

$$\hat{V}_{T_3} = \hat{V}_{T_2} + \hat{V}_\Delta + 2\hat{C}_{T_2\Delta} \quad , \text{ and} \tag{10c}$$

$$\hat{D}_{T_3T_2} = 1 + \hat{D}_{\Delta T_2} \quad . \tag{10d}$$

Similarly if parameter estimates were derived from the original model of equations (8a), (8b), (8c), (8d), and (8e), growth statistics can be obtained by:

$$\hat{D}_{T_3T_2} = \hat{C}_{T_2T_3} \div \hat{V}_{T_2} \quad , \tag{11a}$$

$$\hat{D}_{\Delta T_2} = \hat{D}_{T_3T_2} - 1 \quad , \tag{11b}$$

$$\hat{V}_\Delta = \hat{V}_{T_2} + \hat{V}_{T_3} - 2\hat{C}_{T_2T_3} \tag{11c}$$

$$\hat{C}_{T_2\Delta} = \hat{D}_{\Delta T_2}\hat{V}_\Delta \text{ , and} \tag{11d}$$

$$\hat{\rho}_{T_2\Delta} = \hat{D}_{\Delta T_2}\sqrt{\hat{V}_{T_2} \div \hat{V}_\Delta} \quad . \tag{11e}$$

Following Jöreskog (1971) the parallel test assumption can be tested (given multivariate normality) by comparing the chi-square for the "essentially tau-equivalent" model to that for the "parallel" test model; the difference in chi-square with $df = 2$ is a test of assumptions that $V_{e_{12}} = V_{e_{22}}$ and $V_{e_{13}} = V_{e_{23}}$. Similarly the increase in chi-square from the "congeneric" model to the "essentially tau-equivalent" model ($df = 2$) is a test of the assumptions that $A_{12} = A_{22}$ and $A_{13} = A_{23}$. If the parallel test assumptions are accepted then the population reliability at the initial time may be estimated by $\hat{V}_{T_2} \div (\hat{V}_{T_2} + \hat{V}_{e_{12}})$ and reliability at the final time by $\hat{V}_{T_3} \div (\hat{V}_{T_3} + \hat{V}_{e_{13}})$. The reliability for each test is the square of the corresponding standardized factor loading in the case of "essentially tau-equivalent" or "congeneric" measures. Another statistic of interest in the traditional psychometric literature is the reliability of differences (ρ_Δ) which is defined as the true variance of the differences divided by the variance of the observed differences. In the parallel case the estimated population error variances can be used to obtain $\hat{\rho}_\Delta$ directly:

$$\hat{\rho}_\Delta = \frac{\hat{V}_\Delta}{\hat{V}_\Delta + \hat{V}_{e_{12}} + \hat{V}_{e_{13}}} \quad . \tag{12a}$$

With "essentially tau-equivalent" assumptions no statement is made about equality of error variances so that four reliabilities may be estimated:

$$\hat{\rho}'_\Delta = \frac{\hat{V}_\Delta}{\hat{V}_\Delta + \hat{V}_{e_{12}} + \hat{V}_{e_{13}}} \tag{12b}$$

$$\hat{\rho}''_\Delta = \frac{\hat{V}_\Delta}{\hat{V}_\Delta + \hat{V}_{e_{12}} + \hat{V}_{e_{23}}} \quad , \tag{12c}$$

$$\hat{\rho}'''_\Delta = \frac{\hat{V}_\Delta}{\hat{V}_\Delta + \hat{V}_{e_{22}} + \hat{V}_{e_{13}}} \quad , \tag{12d}$$

$$\hat{\rho}''''_\Delta = \frac{\hat{V}_\Delta}{\hat{V}_\Delta + \hat{V}_{e_{22}} + \hat{V}_{e_{23}}} \quad . \tag{12e}$$

Formulas (12a), (12b), (12c), (12d), and (12e) are based on the assumption that the true scores have the same units as the observed scores, which is not true in the case of congeneric measures. Since the regression of observed on true differences is equal to the regression of observed on true scores (Werts & Linn, 1970a, equation (25)). it is only necessary to standardize this weight with the appropriate variances to obtain the reliability of differences for all cases, e.g., in the congeneric case if $A_{12} = A_{13}$ then

$$\hat{\rho}_{\Delta} = \hat{A}^2_{12} \frac{\hat{V}_{\Delta}}{\hat{V}_{y_{12}} + \hat{V}_{y_{13}} - 2\hat{C}(y_{12}, y_{13})} \tag{12f}$$

where $\hat{V}_{y_{12}}$, $\hat{V}_{y_{13}}$ and $\hat{C}(y_{12}, y_{13})$ are the estimated elements in $\hat{\Sigma}$. This formula uses estimated elements in $\hat{\Sigma}$ which are provided in the computer output for Jöreskog's program (Jöreskog, Gruvaeus, & van Thillo, 1970). The program computes the elements in $\hat{\Sigma}$ from the estimates for the underlying parameters, e.g., $\hat{C}(y_{12}, y_{13}) = \hat{A}_{12}\hat{A}_{13}\hat{C}_{T_2T_3}$. This model (all measurement errors independent) may be used to clarify traditional procedures for obtaining growth statistics. For example, consider the case in which one initial and one final test is given. A common procedure is to obtain split half reliabilities at each time and use these to correct for attenuation. If y_{12} and y_{22} are the initial split halves and y_{13} and y_{23} the final split halves, this case corresponds exactly to the parallel measure case analyzed above. The difference from the traditional procedure is that the complete variance-covariance matrix for the split halves is computed and used in the analysis. As shown above, the "parallel" and "essentially tau-equivalent" assumptions can be tested against the congeneric model and the congeneric model is overidentified. From this perspective the traditional procedure neglects useful information about correlations among split halves and thereby loses the possibility of rejecting the model because of poor fit to the data and of analyzing the data making only congeneric test assumptions. To understand the connection with the traditional formula it is of interest to standardize $\hat{\underset{\sim}{\Sigma}}$ into a correlation matrix (correlations generated by the model are indicated by symbol ρ) and to show the relationships to standardized model parameters (denoted by asterisk):

$$\rho(y_{12}, y_{13}) = \hat{A}^*_{12}\hat{\rho}_{T_2T_3}\hat{A}^*_{13} \tag{13a}$$

$$\rho(y_{12}, y_{23}) = \hat{A}^*_{12}\hat{\rho}_{T_2T_3}\hat{A}^*_{23} \tag{13b}$$

$$\rho(y_{22}, y_{13}) = \hat{A}^*_{22}\hat{\rho}_{T_2T_3}\hat{A}^*_{13} \tag{13c}$$

$$\rho(y_{22}, y_{23}) = \hat{A}^*_{22}\hat{\rho}_{T_2T_3}\hat{A}^*_{23} \tag{13d}$$

$$\rho(y_{12}, y_{22}) = \hat{A}^*_{12}\hat{A}^*_{22} \tag{13e}$$

$$\rho(y_{13}, y_{23}) = \hat{A}^*_{13}\hat{A}^*_{23} \quad . \tag{13f}$$

If parallel test assumptions are valid then $\hat{A}^*_{12} = \hat{A}^*_{22}$ and $\hat{A}^*_{13} = \hat{A}^*_{23}$, in which case equations (13a), (13b), (13c), and (13d) are identical and should be recognized as the traditional correction for attenuation, except that the correlations are drawn from $\hat{\underset{\sim}{\Sigma}}$ rather than from the observed correlation matrix $\underset{\sim}{S}$. Equations (13e) and (13f), under parallel test assumptions, are simply the assumption that the reliability defined as the squared correlation (i.e., A^*_{12} or A^*_{13}) of the observed with the true score is equal to the correlation between two parallel tests, but again the correlations are drawn from $\hat{\Sigma}$ not from S. What these equations show is that it is not necessary for the reliabilities of the split halves to be equal in order to identify the unattenuated correlation $\hat{\rho}_{T_2T_3}$ given uncorrelated errors. If the estimates of the elements in $\hat{\Sigma}$ for the parallel case are examined it will be found that because of the structural specifications: $\hat{V}_{y_{12}} = \hat{V}_{y_{22}}$, $\hat{V}_{y_{13}} = \hat{V}_{y_{23}}$, $\hat{C}(y_{12}, y_{13}) = \hat{C}(y_{13}, y_{23}) = \hat{C}(y_{22}, y_{13}) = \hat{C}(y_{22}, y_{23})$, $\hat{C}(y_{12}, y_{22}) = \hat{V}_{T_2}$, $\hat{C}(y_{13}, y_{23}) = \hat{V}_{T_3}$ and $\hat{C}(y_{12}, y_{13}) = \hat{C}(y_{22}, y_{23}) = \hat{C}_{T_2T_3}$. Translating the equation for the reliability of differences into the elements of $\hat{\Sigma}$:

$$\hat{\rho}_{\Delta} = \frac{\hat{C}(y_{12}, y_{22}) + \hat{C}(y_{13}, y_{23}) - 2\hat{C}(y_{12}, y_{13})}{\hat{V}_{y_{12}} + \hat{V}_{y_{13}} - 2\hat{C}(y_{12}, y_{13})} \tag{14a}$$

or

$$\hat{\rho}_{\Delta} = \frac{\hat{V}_{y_{12}}\hat{\rho}(y_{12}, y_{22}) + \hat{V}_{y_{13}}\hat{\rho}(y_{13}, y_{23}) - 2\hat{\rho}(y_{12}, y_{13})\sqrt{\hat{V}_{y_{12}}\hat{V}_{y_{13}}}}{\hat{V}_{y_{12}} + \hat{V}_{y_{13}} - 2\hat{\rho}(y_{12}, y_{13})\sqrt{\hat{V}_{y_{12}}\hat{V}_{y_{13}}}}. \tag{14b}$$

Equation (14b) should be recognized as the traditional formula for the reliability of differences, noting however that the estimates are drawn from $\hat{\Sigma}$, not from the observed matrix S. The essentially tau-equivalent case differs from the parallel case in that the corresponding variances in $\hat{\Sigma}$ are not required to be equal, however the covariances between independent measures of different traits are still equal to the covariances between the corresponding traits factors. This means that formula (14a) could be used for any pair of tau-equivalent tests over time. For congeneric measures the formula involves the pairs of measures which have the same units over time, e.g., if $A_{12} = A_{13}$ then equation (12f) may be translated into

$$\hat{\rho}_{\Delta} = \frac{\hat{V}_{y_{12}} (\hat{A}^*_{12})^2 + \hat{V}_{y_{13}} (\hat{A}^*_{13})^2 - 2\hat{A}^*_{12}\hat{A}^*_{13}\hat{\rho}_{T_2T_3} \sqrt{\hat{V}_{y_{12}}\hat{V}_{y_{13}}}}{\hat{V}_{y_{12}} + \hat{V}_{y_{13}} - 2\hat{\rho}(y_{12}, y_{13}) \sqrt{\hat{V}_{y_{12}}\hat{V}_{y_{13}}}} \quad . \quad (14c)$$

Equation (14c) is the reliability of differences formula given by Werts and Linn (1970a, equation (26)) for the case of correlated errors over time for the pair of measurements on the same scale, i.e., the Werts and Linn formula is also appropriate to the independent error case when applied to the elements of $\hat{\underset{\sim}{\Sigma}}$ rather than $\underset{\sim}{S}$. If formula (14c) applies to correlated errors using congeneric measures then it may be specialized for the parallel measures case, e.g., if y_{12} and y_{13} have nonindependent errors and y_{12} and y_{23} have independent errors:

(a) $\hat{A}^*_{13} = \hat{A}^*_{23}$, by parallel test assumptions, therefore $\hat{A}^*_{12}\hat{A}^*_{13}\hat{\rho}_{T_2T_3} = \hat{A}^*_{12}\hat{A}^*_{23}\hat{\rho}_{T_2T_3}$,

(b) but $\hat{\rho}_{(y_{12}, y_{23})} = \hat{A}^*_{12}\hat{A}^*_{23}\hat{\rho}_{T_2T_3}$.

Since

$$\hat{A}^*_{12} = \sqrt{\hat{\rho}(y_{12}, y_{22})}\ , \quad \hat{A}^*_{13} = \sqrt{\hat{\rho}(y_{13}, y_{23})}\ , \quad \hat{V}_{y_{12}} = \hat{V}_{y_{22}}\ , \quad \hat{V}_{y_{13}} = \hat{V}_{y_{23}}\ ,$$

equation (14c) becomes

$$\hat{\rho}_{\Delta} = \frac{\hat{V}_{y_{12}}\hat{\rho}(y_{12}, y_{22}) + \hat{V}_{y_{13}}\hat{\rho}(y_{13}, y_{23}) - 2\hat{\rho}(y_{12}, y_{23}) \sqrt{\hat{V}_{12}\hat{V}_{13}}}{\hat{V}_{y_{12}} + \hat{V}_{y_{13}} - 2\hat{\rho}(y_{12}, y_{13}) \sqrt{\hat{V}_{12}\hat{V}_{13}}} \quad . \quad (14d)$$

Equation (14d) is the formula for the reliability of differences for "linked" (i.e., correlated errors) parallel test measures given by Cronbach and Furby (1970, equation (6)), which can be seen to be the parallel measure specialization of the Werts-Linn equation for nonindependent congeneric measures.

Similarly from equations (11a), (11b), (11c), (11d), and (11e) it follows that the estimated correlation of status with gain is:

$$\hat{\rho}_{T_2\Delta} = \frac{\hat{C}_{T_3T_2} - \hat{V}_{T_2}}{\sqrt{\hat{V}_{T_2}(\hat{V}_{T_2} + \hat{V}_{T_3} - 2\hat{C}_{T_2T_3})}} \quad . \tag{15a}$$

In the congeneric case with $A_{12} = A_{13}$, this may be transformed into

$$\hat{\rho}_{T_2\Delta} = \frac{\hat{\rho}_{T_2T_3}\, \hat{A}^*_{13}\sqrt{\hat{V}_{y_{13}}} - \hat{A}^*_{12}\sqrt{\hat{V}_{y_{12}}}}{\sqrt{\hat{A}^{*2}_{12}\hat{V}_{y_{12}} + \hat{A}^{*2}_{13}\hat{V}_{y_{13}} - 2\hat{\rho}_{T_2T_3}\hat{A}^*_{12}\hat{A}^*_{13}\sqrt{\hat{V}_{y_{12}}\hat{V}_{y_{13}}}}} \quad . \tag{15b}$$

Formula (15b) is the correlation of status with gain given by Werts and Linn (1970a, equation (28)) for the case of congeneric measures and correlated errors, i.e., the formula applies also to the independent error case. In the case of parallel independent measures $\hat{\rho}_{T_2T_3} = \hat{\rho}(y_{12}, y_{13}) \div \sqrt{\hat{\rho}(y_{12}, y_{22})\hat{\rho}(y_{13}, y_{23})}$ which when substituted into formula (15b) yields the traditional formula for the correlation of status with gain as applied to the elements of $\hat{\underset{\sim}{\Sigma}}$:

$$\hat{\rho}_{T_2\Delta} = \frac{\hat{\rho}(y_{12}, y_{13})\sqrt{\hat{V}_{y_{13}}} - \hat{\rho}(y_{12}, y_{22})\sqrt{\hat{V}_{y_{12}}}}{\sqrt{\hat{\rho}(y_{12}, y_{22}}\sqrt{\hat{\rho}(y_{12}, y_{22})\,\hat{V}_{y_{12}} + \hat{\rho}(y_{13}, y_{23})\,\hat{V}_{y_{13}} - 2\hat{\rho}(y_{12}, y_{13})\sqrt{\hat{V}_{y_{12}}\hat{V}_{y_{13}}}}} \quad . \tag{15c}$$

Our purpose in demonstrating relationships to traditional formulations is purely heuristic, since Jöreskog's program yields estimates of model parameters given the structural assumptions specified by the investigator, i.e., the traditional formulas apply to the elements of $\hat{\underset{\sim}{\Sigma}}$ which are not directly observable but which are estimated as a function of the parameter estimates.

Traditional psychometric approaches have dealt with models which are just identified which means that models which exactly reproduce the observed variance-covariance matrix can be employed (i.e., $\underset{\sim}{S} = \hat{\underset{\sim}{\Sigma}}$). The limitation in this approach is that overidentification is necessary if the fit of the model to the data is to be tested.

In this paragraph we propose to use our model to specify the conditions implicit in Cronbach's (1960, pp. 136-139) discussion of coefficients of "stability" and "equivalence." Cronbach uses an example in which two forms of the Mechanical Reasoning Test of the DAT were used, the same forms being used for test and retest purposes. When the same form is repeated, the test-retest correlation is higher than the test-retest correlation between different forms, suggesting the presence of "long-lasting test-specific" factors. The implication is that the errors of measurement for the same test repeated are not independent. Assuming that both forms were repeated and errors of measurement independent for different forms, the model for parallel measures is of the form:

$$\underset{\sim}{y} = \underset{\sim}{\mu} + \underset{\sim}{\Lambda}\underset{\sim}{T} \ , \tag{16a}$$

where

$$\underset{\sim}{y} = (y_{12}, y_{22}, y_{13}, y_{23}) \tag{16b}$$

where y_{12} and y_{13} are the same test as are y_{22} and y_{23} .

$$\underset{\sim}{T} = (T_2, T_3, e_{12}, e_{22}, e_{13}, e_{23}) \quad , \tag{16c}$$

$$\underset{\sim}{\Lambda} = \begin{bmatrix} 1 & 0 & 1 & 0 & 0 & 0 \\ 1 & 0 & 0 & 1 & 0 & 0 \\ 0 & 1 & 0 & 0 & 1 & 0 \\ 0 & 1 & 0 & 0 & 0 & 1 \end{bmatrix} , \qquad (16d)$$

and

$$\underset{\sim}{\Phi} = \begin{bmatrix} V_{T_2} & & & \text{Symmetric} & & \\ C_{T_2T_3} & V_{T_3} & & & & \\ 0 & 0 & V_{E_{12}} & & & \\ 0 & 0 & 0 & V_{e_{22}} & & \\ 0 & 0 & C_{e_{12}e_{13}} & 0 & V_{e_{13}} & \\ 0 & 0 & 0 & C_{e_{22}e_{23}} & 0 & V_{e_{23}} \end{bmatrix} , \qquad (16e)$$

where $V_{e_{12}} = V_{e_{22}}$, $V_{e_{13}} = V_{e_{23}}$.

The model of (16a) is the special case of factor analysis in which the residual factors are treated as latent factors. Examination of $\underset{\sim}{\Phi}$ shows that the same test errors of measurement are nonindependent, i.e., $C_{e_{12}e_{22}}$ and $C_{e_{22}e_{23}} \neq 0$. All parameters are identifiable and df = 3 (10 distinct elements in $\underset{\sim}{\Sigma}$ less 7 free and constrained parameters). Essentially tau-equivalent assumptions would still have provided identification but with only one overidentifying restriction (since $V_{e_{12}} \neq V_{e_{22}}$, $V_{e_{13}} \neq V_{e_{23}}$) . An interesting case occurs with congeneric assumptions in which case the model is underidentified; however, the unattenuated trait correlation $\rho_{T_2T_3}$ is just identified

$[\hat{\rho}^2_{T_2T_3} = \hat{C}(y_{12}, y_{23})\hat{C}(y_{22}, y_{13}) \div \hat{C}(y_{12}, y_{22})\hat{C}(y_{13}, y_{23})]$. Identification may be achieved with the congeneric model by repeating only one test (assuming $A_{12} = A_{13}$) and using different method measures for y_{22} and y_{23} in which case the model is:

$$\underset{\sim}{\Lambda} = \begin{bmatrix} A_{12} & 0 & 1 & 0 & 0 & 0 \\ A_{22} & 0 & 0 & 1 & 0 & 0 \\ 0 & A_{13} & 0 & 0 & 1 & 0 \\ 0 & A_{23} & 0 & 0 & 0 & 1 \end{bmatrix} , \tag{16f}$$

where $A_{12} = A_{13}$ by assumption, and

$$\Phi = \begin{bmatrix} 1 & & & & & \\ C_{T_2T_3} & V_{T_3} & & & \text{Symmetric} & \\ 0 & 0 & V_{e_{12}} & & & \\ 0 & 0 & 0 & V_{e_{22}} & & \\ 0 & 0 & C_{e_{12}e_{13}} & 0 & V_{e_{13}} & \\ 0 & 0 & 0 & 0 & 0 & V_{e_{23}} \end{bmatrix} . \tag{16g}$$

This model is just identified (10 distinct elements in $\underset{\sim}{\Sigma}$ less 10 parameters to be estimated). Let us return to Cronbach's example in which there are Forms A (y_{12}) and B (y_{22}) initially and retests on Forms A (y_{23}) and B (y_{23}) three years later. Cronbach partitions the variance using the immediate and retest correlations among forms (assumed parallel) which in our model corresponds to the elements of $\underset{\sim}{\Sigma}$. We may translate Cronbach's partitioning procedure into functions of the model parameters in equations (16a), (16b), (16c), (16d), and (16e) as follows:

1. "Lasting General Variance" = $\rho(y_{12}, y_{23}) = A^*_{12}\rho(T_2, T_3)A^*_{23}$ which according to the model equals $\rho(y_{22}, y_{13}) = A^*_{22}\rho(T_2, T_3)A^*_{13}$.

2. "Temporary General Variance" = $\rho(y_{12}, y_{22}) - \rho(y_{12}, y_{23}) = A^*_{12}A^*_{22} - A^*_{12}\rho(T_2, T_2)A^*_{23}$ which according to the model equals $\rho(y_{12}, y_{22}) - \rho(y_{22}, y_{13}) = A^*_{12}A^*_{22} - A^*_{22}\rho(T_2, T_3)A^*_{13}$. In principle there is a different "Temporary General Variance" for the end time $\rho(y_{13}, y_{23}) - \rho(y_{12}, y_{23}) = A^*_{13}A^*_{23} - A^*_{12}\rho(T_2, T_3)A^*_{23}$ which equals $\rho(y_{13}, y_{23}) - \rho(y_{22}, y_{13}) = A^*_{13}A^*_{23} - A^*_{22}\rho(T_2, T_3)A^*_{13}$.

3. "Lasting Specific Variance" for Form A $\rho(y_{12}, y_{13}) - \rho(y_{12}, y_{23}) = \rho(y_{12}, y_{13}) - \rho(y_{22}, y_{13}) = \sqrt{1 - (A^*_{12})^2}\, \rho(e_{12}, e_{13}) \sqrt{1 - (A^*_{13})^2}$ and for Form B $\rho(y_{22}, y_{23}) - \rho(y_{12}, y_{23}) = \rho(y_{22}, y_{23}) - \rho(y_{22}, y_{13}) = \sqrt{1 - (A^*_{22})^2}\, \rho(e_{22}, e_{23}) \sqrt{1 - (A^*_{23})^2}$.

4. "Temporary Specific Variance" $[1 - \rho(y_{12}, y_{22})] - [\rho(y_{12}, y_{13}) - \rho(y_{12}, y_{23})] = [1 - \rho(y_{12}, y_{22})] - [\rho(y_{12}, y_{13}) - \rho(y_{22}, y_{13})] = 1 - A^*_{12}A^*_{22} - \sqrt{1 - (A^*_{12})^2}\, \rho(e_{12}, e_{13}) \sqrt{1 - (A^*_{13})^2}$ for the correlations used by Cronbach, but in principle there are three other temporary specific variances $1 - A^*_{12}A^*_{22} - \sqrt{1 - (A^*_{22})^2}\, \rho(e_{22}, e_{23}) \sqrt{1 - (A^*_{23})^2}$, $1 - A^*_{13}A^*_{23} - \sqrt{1 - (A^*_{12})^2}\, \rho(e_{12}, e_{13}) \sqrt{1 - (A^*_{13})^2}$, and $1 - A^*_{13}A^*_{23} - \sqrt{1 - (A^*_{22})^2}\, \rho(e_{22}, e_{23}) \sqrt{1 - (A^*_{23})^2}$.

It can be seen that Cronbach's procedure for partitioning of variance involves complicated functions of the model parameters. Not only is it simpler to analyze observed correlations in terms of a set of structural parameters, but it allows for analysis of overidentified models. Further light can be shed on the assumptions implicit in the model of (16a), (16b), (16c), (16d), and (16e) by asking what variables account for the correlated errors. Assuming that a single factor (M_1) underlies the correlation for Form A and another factor (M_2) for Form B the model becomes:

$$\underset{\sim}{y} = \underset{\sim}{\mu} + \underset{\sim}{\Lambda}\underset{\sim}{T} + \underset{\sim}{e}' \quad , \tag{17a}$$

where

$$\underset{\sim}{y} = (y_{12}, y_{22}, y_{13}, y_{23}) \quad , \tag{17b}$$

$$\underset{\sim}{T} = (T_2, T_3, M_1, M_2) \quad , \tag{17c}$$

$$\underset{\sim}{e}' = (e'_{12}, e'_{22}, e'_{13}, e'_{23}) \quad , \tag{17d}$$

$$\underset{\sim}{\Lambda} = \begin{bmatrix} 1 & 0 & B_{12} & 0 \\ 1 & 0 & 0 & B_{22} \\ 0 & 1 & B_{13} & 0 \\ 0 & 1 & 0 & B_{23} \end{bmatrix} , \tag{17e}$$

and

$$\underset{\sim}{\Phi} = \begin{bmatrix} V_{T_2} & & & \\ C_{T_2T_3} & V_{T_3} & & \\ 0 & 0 & 1 & \\ 0 & 0 & 0 & 1 \end{bmatrix} . \tag{17f}$$

Analysis of the identification problem shows that B_{12}, B_{22}, B_{13}, and B_{23} are not separately identifiable; only the products $(B_{12} B_{13})$ and $(B_{22} B_{23})$ are identified. This means that in Jöreskog's program we may arbitrarily set $B_{12} = B_{13}$ and $B_{22} = B_{23}$ without disturbing the estimation for other parameters. Assuming $B_{12} = B_{13}$ and $B_{22} = B_{23}$, this model is a simple transformation of (16a), (16b), (16c), (16d), and (16e) under essentially tau-equivalent assumptions, that is, $V_{e_{12}} \neq V_{e_{22}}$, $V_{e_{13}} \neq V_{e_{23}}$ in equation (16e). In particular it can be seen that it must be assumed that M_1 and M_2 are uncorrelated. It is possible to deal with oblique true and method factors but usually more different method measures are required as in our 3 trait x 3 method example in Section I.

When methods of measuring a trait are made as different as possible, it is usually the case that the units of measurement are different, which means that congeneric rather than essentially tau-equivalent or parallel assumptions are appropriate. Werts and Linn (1970a) consider growth models based on congeneric measures, e.g., in one case they use three congeneric measures of T_2 and two congeneric measures of T_3, allowing for same test correlated errors over time. This model is overidentified, but no attempt was made to deal with this complication. Phrasing this problem in terms of Jöreskog's general model:

$$\underset{\sim}{y} = \underset{\sim}{\mu} + \underset{\sim}{\Lambda}\underset{\sim}{T} + \underset{\sim}{e} \tag{18a}$$

$$\underset{\sim}{y} = (y_{12}, y_{22}, y_{32}, y_{13}, y_{23}) \tag{18b}$$

where y_{12} and y_{13} are linked as are y_{22} and y_{23}.

$$\underset{\sim}{T} = (T_2, T_3, M_1, M_2) \tag{18c}$$

$$\Lambda = \begin{bmatrix} A_{12} & 0 & B_{12} & 0 \\ A_{22} & 0 & 0 & B_{22} \\ A_{32} & 0 & 0 & 0 \\ 0 & A_{13} & B_{13} & 0 \\ 0 & A_{23} & 0 & B_{23} \end{bmatrix} \tag{18d}$$

$$\Phi = \begin{bmatrix} 1 & & & \\ C_{T_2T_3} & V_{T_3} & & \\ 0 & 0 & 1 & \\ 0 & 0 & 0 & 1 \end{bmatrix} . \tag{18e}$$

Assuming that $A_{12} = A_{13}$, $A_{22} = A_{23}$ and for convenience that $B_{12} = B_{13}$, $B_{22} = B_{23}$, this model has four overidentifying restrictions (15 distinct elements in $\underset{\sim}{\Sigma}$ less 11 parameters to be estimated). Werts and Linn give two formulas (1970a, p. 198, equations (28) and (29)) for estimating the correlation of status with gain involving observed correlations and variances whereas Jöreskog's approach generates a single estimate by equation (15a). In essence Werts and Linn dealt with the elements of the observed variance-covariance matrix $\underset{\sim}{S}$ which may yield inconsistent estimates of $\rho_{T_2\Delta}$ whereas such inconsistency cannot arise with respect to the elements in $\hat{\underset{\sim}{\Sigma}}$. Jöreskog has an unpublished operating program for estimating factor scores within the confirmatory factor analysis model (Jöreskog, 1971). As Cronbach and Furby (1970) note, however, there is seldom need for such estimates.

Relationship to Factor Analysis

A common practice in the factor analysis of growth data is to compare standardized factor loadings at one time to the loadings for the same set of measures at a later time. If the pattern of loadings remains constant over time the inference is drawn that the factors are measuring essentially the same dimension at different times. For example we might have three measures of T_2 at time 1 with factor loadings $A^*_{12} = .30$, $A^*_{22} = .40$, and $A^*_{32} = .50$ and identical loadings on T_3 when these measures are repeated at time 2, i.e., $A^*_{13} = .30$, $A^*_{23} = .40$, and $A^*_{33} = .50$. For heuristic purposes let us suppose that the repetition of tests did not result in methods factors and that the true variance increased from $V_{T_2} = 1.0$ to $V_{T_3} = 1.5$ over time and $C_{T_2T_3} = 1.2$. It may be immediately inferred that the error variances for all tests increased over time since the test reliabilities (in this model the squared factor loadings) remained constant and the true variance increased. However, Wiley and Wiley (1970) have persuasively argued that it is more likely that error variances are a test characteristic which is likely to remain constant over time. If this is so, then an increase in true variance along the same dimension will necessarily mean that the reliabilities of the tests will increase over time, i.e., the standardized factor loadings will increase. In the same fashion it may be deduced that if for any given test over time the unstandardized regression weights $(A_{i2} = A_{i3})$ and the error variances $(V_{e_{i2}} = V_{e_{i3}})$ are equal, then in general the standardized factor loadings (A^*_{ij}) are not proportional from one time to another. We conclude that comparison of standardized factor loading patterns over time provides no logical base for any conclusions about whether pretests and posttests are measuring the same variable. It appears to us that such an assumption, which in this model is equivalent to equality of unstandardized regression weights over time (e.g., $A_{12} = A_{13}$) , is basically not testable within the framework of this model. It would seem better not to make dubious assumptions that either the reliability or the error variance are relatively constant (over time) test characteristics, but to build models and gather requisite information such that these model parameters are identified.

While it is not possible to test the assumption that $A_{12} = A_{13}$, it is quite possible for this assumption to be incompatible with the assumption that $A_{22} = A_{23}$. The ratio of V_{T_3} to V_{T_2} resulting from $A_{12} = A_{13}$ may differ from the ratio resulting from $A_{22} = A_{23}$. This may be tested by the increase in χ^2 (df = 1) resulting from the addition of $A_{22} = A_{23}$ to the model in which $A_{12} = A_{13}$. Within the framework of this model, if it is true that the corresponding pairs of tests over time in fact have the same units, then the scaling of V_{T_3} to V_{T_2} should be the same for each pair.

The finding that the data are consistent with the hypothesis that $A_{12} = A_{13}$ and $A_{22} = A_{23}$ does not necessarily imply that the units of measurement for the corresponding pairs of tests over time are the same since it is quite possible for the scaling to be erroneous for both pairs of tests but in the same way. If the data are inconsistent with the hypothesis that $A_{12} = A_{13}$ and $A_{22} = A_{23}$ we could conclude that the units over time are not the same for both sets of tests, but it is still possible that the units are the same for one of the sets over time. Even if it could be shown that $A_{12} = A_{13}$, this would only be evidence consistent with, not proof of, the hypothesis that the scales are measuring the same process over time.

Determinants of Growth

Werts and Linn (1970b) have considered the problem of making inferences about the determinants in a linear model. The Werts-Linn formulation was based on classical true score assumptions, i.e., no provision was made for methods factors. For heuristic purposes let us reconsider the problem of growth determinants, formulating the three trait, three method model in terms of growth $(T_3 = T_2 + \Delta)$:

$$\underset{\sim}{T} = (T_1, T_2, \Delta, M_1, M_2, M_3) \tag{19a}$$

$$\underset{\sim}{B} = \begin{bmatrix} A_{11} & 0 & 0 & B_{11} & 0 & 0 \\ A_{21} & 0 & 0 & 0 & B_{21} & 0 \\ A_{31} & 0 & 0 & 0 & 0 & B_{31} \\ 0 & A_{12} & 0 & B_{12} & 0 & 0 \\ 0 & A_{22} & 0 & 0 & B_{22} & 0 \\ 0 & A_{32} & 0 & 0 & 0 & B_{32} \\ 0 & A_{12} & A_{12} & B_{13} & 0 & 0 \\ 0 & A_{22} & A_{22} & 0 & B_{23} & 0 \\ 0 & A_{32} & A_{32} & 0 & 0 & B_{33} \end{bmatrix} \tag{19b}$$

$$\Phi = \begin{bmatrix} 1 & & & & & \\ C_{T_1T_2} & 1 & & & & \\ C_{T_1\Delta} & C_{T_2\Delta} & V_\Delta & & & \\ C_{T_1M_1} & C_{T_2M_1} & C_{\Delta M_1} & 1 & & \\ C_{T_1M_2} & C_{T_2M_2} & C_{\Delta M_2} & C_{M_1M_2} & 1 & \\ C_{T_1M_3} & C_{T_2M_3} & C_{\Delta M_3} & C_{M_1M_3} & C_{M_2M_3} & 1 \end{bmatrix} . \qquad (19c)$$

It should be noted that although this formulation does not directly involve the parameters of the underlying growth model $\Delta = D_{\Delta T_1 \cdot T_2} T_1 + D_{\Delta T_2 \cdot T_1} T_2 + \xi$, however, the regression weights are:

$$D_{\Delta T_1 \cdot T_2} = \frac{C_{T_1\Delta} - C_{T_2\Delta}C_{T_1T_2}}{1 - C^2_{T_1T_2}} , \qquad (19d)$$

and

$$D_{\Delta T_2 \cdot T_1} = \frac{C_{T_2\Delta} - C_{T_1\Delta}C_{T_1T_2}}{1 - C^2_{T_1T_2}} . \qquad (19e)$$

Traditional test theorists (e.g., Bloom, 1964; Thorndike, 1966) have been very concerned with and have drawn substantive inferences about the determinants of growth from the correlation of status with gain, usually corrected for "attenuation." However, as detailed by Werts and Linn (1970b), in a linear structural model prime interest is in the model parameters $D_{\Delta T_1 \cdot T_2}$ and $D_{\Delta T_2 \cdot T_1}$ since if either one is zero the inference will be drawn that the corresponding variable does not directly influence gain. Except in the case in which initial status is uncorrelated with all determinants of growth, knowledge of the correlation of status with gain, $\rho_{T_2\Delta}$, does not allow us to draw inferences about model parameters. It is quite possible for $\rho_{T_2\Delta}$ to be completely spurious due to a common antecedent influence or it is quite possible for $\rho_{T_2\Delta}$ to be zero without implying that $D_{\Delta T_1 \cdot T_2}$ or $D_{\Delta T_2 \cdot T_1}$ be zero. For this reason we question Thorndike's (1966, p. 124) interpretation: "In considerable part, the factors that produce gains during a specified time span appear to be different from those that produced the level of competence exhibited at the beginning of the period." Our objection is that Thorndike's conclusion was made from the correlation of status with gain, without specifically introducing into the analysis any presumed determinants of growth. In a linear structural model the total association of initial status with growth is an insufficient basis for drawing inferences about the various possible determinants of growth.

Discussion

The variety of test response tendencies covered by the rubric "methods factors" appear to be an almost universal complication in sociopsychological growth studies. Even though in principle the multitrait-multimethod model presented in this paper provides for "methods factors," it does not follow that this model does in fact provide a better simulation of reality than previous models which have typically ignored methods factors by assuming independent errors of measurement. It may be expected that our procedure will typically yield different parameter estimates (e.g., correlation of status with gain) than previous procedures, but what has been learned about growth and its determinants thereby? What is learned about reality from the overwhelming concern of the factor analyst with statistical fit? There is no guarantee that the best fitting model yields substantively meaningful results (e.g., Werts, Jöreskog, & Linn, in press). Why bother with complicated structural models involving unmeasured variables when it is likely that a simple regression equation involving only measured variables will provide the best prediction of the criterion? From our perspective, if the researcher's basic interest is in reality, then the research must be designed to explore reality, i.e., to offer evidence as to which of the initially plausible alternative hypotheses (models) provides the better simulation. In some cases this may involve a study of the theoretical implications to see what information is necessary to discriminate between the alternative models. In other cases the study may be a continuing one as in the building of models to simulate the national economy, in which case the ability to better predict new yearly data is used to discriminate among models. Our purpose in making these remarks is to heighten the awareness of researchers that parameter estimates, such as the reliability of gain scores, are always made within the framework of a whole set of untested assumptions about the nature of reality. It is misleading to talk about "the correlation of status with gain" since the meaning of this parameter is totally a function of the particular model used to derive the parameter. In most cases in which this type of estimate has been used, no effort has been made to examine the validity or even plausibility of the models underlying these estimates. The linear structural model presented herein is as suspect as any other model and needs to be justified as one of the plausible alternative hypotheses, prior to data analysis.

References

Bloom, B. S. Stability and change in human characteristics. New York: Wiley, 1964.

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

Cronbach, L. J. Essentials of psychological testing. New York: Harper & Brothers, 1960.

Cronbach, L. J., & Furby, L. How we should measure "change"--or should we? Psychological Bulletin, 1970, 74, 68-80.

Fisher, F. M. The identification problem in econometrics. New York: McGraw-Hill, 1966.

Hauser, R. M., & Goldberger, A. S. The treatment of unobservable variables in path analysis. Social Systems Research Institute Workshop Series EME7030, University of Wisconsin, August 1970.

Jöreskog, K. G. A general approach to confirmatory maximum likelihood factor analysis. Psychometrika, 1969, 34, 183-202.

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

Jöreskog, K. G. A general method for estimating a linear structural equation system. Research Bulletin 70-54. Princeton, N. J.: Educational Testing Service, September 1970. (b)

Jöreskog, K. G. Statistical analysis of sets of congeneric tests. Psychometrika, 1971, 36, in press.

Jöreskog, K. G., Gruvaeus, G. T., & van Thillo, M. ACOVS--a general computer program for analysis of covariance structures. Research Bulletin 70-15. Princeton, N. J.: Educational Testing Service, February 1970.

Lord, F. M., & Novick, M. R. Statistical theories of mental test scores. New York: Addison-Wesley, 1968.

Thorndike, R. L. Intellectual status and intellectual growth. Journal of Educational Psychology, 1966, 57, 121-127.

Werts, C. E., & Linn, R. L. Path analysis: psychological examples. Psychological Bulletin, 1970, 74, 193-212. (a)

Werts, C. E., & Linn, R. L. A general linear model for studying growth. Psychological Bulletin, 1970, 73, 17-22. (b)

Werts, C. E., Jöreskog, K. G., & Linn, R. L. Comment on "the estimation of measurement error in panel data." American Sociological Review, in press.

Wiley, D. F., & Wiley, J. A. The estimation of measurement error in panel data. American Sociological Review, 1970, 35, 112-117.