

Longitudinal Models of Reliability and Validity: A Latent Curve Approach

John Tisak and Marie S. Tisak

Bowling Green State University

The concepts of reliability and validity and their associated coefficients typically have been restricted to a single measurement occasion. This paper describes dynamic generalizations of reliability and validity that will incorporate longitudinal or developmental models, using latent curve analysis. Initially a latent curve model is formulated to depict change. This longitudinal model is then incorporated into the classical definitions of reliability and validity. This approach permits the separa-

tion of constancy or change from the indexes of reliability and validity. Statistical estimation and hypothesis testing may be achieved using standard structural equations modeling computer programs. These longitudinal models of reliability and validity are demonstrated on sociological and psychological data. *Index terms: concurrent validity, dynamic models, dynamic true score, latent curve analysis, latent trajectory, predictive validity, reliability, validity.*

The measurement notions of reliability and validity have endured and are of great practical and theoretical importance in applications of psychological measurement (Lord & Novick, 1968; Meier, 1993). However, significant advances have been made in longitudinal modeling and research (Collins & Horn, 1991; Menard, 1991; von Eye & Clogg, 1994) and many longitudinal datasets are currently available (Verdonik & Sherrod, 1984). Unfortunately, the concepts of reliability and validity and their respective coefficients typically have been static; that is, their definitions have been based on observations made at a single occasion or, in the case of test-retest reliability, a single coefficient for two measurement occasions. This restriction has not been maintained in other disciplines, such as sociology in which alternative dynamic models have been proposed (Wheaton, Muthén, Alwin, & Summers, 1977). Additionally, the issue of a dynamic criterion was recognized and discussed over 30 years ago (Bass, 1962; Fleishman & Fruchter, 1960; Ghiselli & Haire, 1960).

This paper explores dynamic generalizations of reliability and validity; that is, those that extend the classical approach to incorporate longitudinal or developmental models using latent curve analysis (LCA; Meredith & Tisak, 1990). Models are proposed that permit reliability and validity to change as a function of time. Hence, although no additional structure has been imposed, the effects of maturation, context, time, and other potential longitudinal influences are permitted to manifest themselves in changes to the reliability and validity of the instrument. Additionally, constancy and change may be addressed separately from reliability and validity (i.e., they are parameterized separately in the models). Further, the approach is readily generalizable to include different groups or populations. Finally, if there is only a single measurement occasion, these longitudinal models reduce to the classical definitions of reliability and validity.

The subsequent developments make use of the standard definitions of the coefficients of reliability and validity (Lord & Novick, 1968). The reliability coefficient is defined as the ratio of true-score variance (σ_T^2) to observed-score variance (σ_X^2):

$$\rho_R = \sigma_T^2 / \sigma_X^2. \quad (1)$$

The validity coefficient is defined as the absolute value of the correlation coefficient between two measures, X and Y :

$$\rho_v = |\rho_{XY}|. \quad (2)$$

The argument that the best way to study longitudinal change is to first formulate a model for change or development was made cogently by Rogosa, Brandt, & Zimowski (1982) and Rogosa & Willett (1985). In keeping with this postulate, a model for change or development is presented in which the longitudinal function may either be linearized (e.g., polynomial growth) or approximately linearized, as in a Maclaurin or Taylor series expansion (Protter & Morrey, 1970). The linearization approach was advocated first by Tucker (1958), Rao (1958), and Meredith in an appendix to Scher, Young, & Meredith (1960) and later extended to a latent variable model by Meredith & Tisak (1990). Specifically, time series measures may be decomposed into latent basis curves and measurement error; the latent curves may be used to depict change. Additionally, the approach is general enough to include polynomial or orthogonal polynomial growth and some spline functions as special cases (Tisak & Meredith, 1990). Finally from this perspective, the concept of reliability (and validity) may be disentangled from the notions of constancy and change. In particular, measurements are decomposed into separate sets of parameters that represent reliability and the function of change.

The procedure that is developed below has several advantages over static reliability and validity coefficients (Equations 1 and 2, respectively) when persons are measured repeatedly. The dynamic models ameliorate some of the problems with the classical model; for example, multiple measures are not needed at a time point as required for internal consistency estimates, and refinements can be made with test-retest estimates, as shown below. Further, this new approach provides a unified approach to the representation of constancy/change and reliability/validity (i.e., parameters are estimated and tested simultaneously). Lastly, when there is change, the method generalizes the concept of true score to a dynamic true score or a true score model that includes change (i.e., true score that is a function of time). In keeping with the longitudinal nature of this approach, the dynamic true score will be referred to as the *latent trajectory*.

Formulation and Estimation

Let the observations or persons be indexed by $i \in \{1, 2, \dots, N\}$ and the time of measurement (occasion) be indexed by s and t . For the i th person at the s th time, let the measured variable, $x_i(s)$, be represented or approximated by

$$x_i(s) = \sum_{j=1}^{r_1} g_j(s)v_{ij} + e_i(s). \quad (3)$$

Similarly, at the t th time, let the measured variable, $y_i(t)$, be given by

$$y_i(t) = \sum_{k=1}^{r_0} h_k(t)w_{ik} + f_i(t). \quad (4)$$

Note that times s and t are discrete (i.e., $s \in \{s_1, s_2, \dots, s_p\}$ and $t \in \{t_1, t_2, \dots, t_q\}$) and that these sets of time points may or may not coincide. Also, note the special cases in which $p = 1$ and/or $q = 1$; that is, this formulation includes those situations in which one or both variables are measured at only one time. r_1 and r_0 are the number of basis curves needed to adequately represent the latent trajectories in Equations 3 and 4, respectively; for parsimony, it is desirable that these limits are small (e.g., 1, 2, or 3). Further, in this situation, $x_i(s)$ and $y_i(t)$ are the observed or manifest variables that are assessed across time; $g_j(s)$ and $h_k(t)$ are the latent basis curves; v_{ij} and w_{ik} are the person saliences or weights that the i th person attaches to the j th and k th curves, respectively; and $e_i(s)$ and $f_i(t)$ are the errors or residuals.

To illustrate, consider the case in which growth is linear (Rogosa et al., 1982); that is, $r_1 = 2$, $g_1(s) \equiv 1$,

$g_2(s) = s$, $v_{11} = a_i$, and $v_{12} = b_i$; then Equation 3 becomes

$$x_i(s) = a_i + b_i s + e_i(s). \quad (5)$$

For this situation, the person parameters a_i and b_i represent initial status and rate of change, respectively, for the i th person, and s represents time.

With the latent curve representation of Equations 3 and 4 and with standard assumptions of measurement errors (i.e., they are uncorrelated with each other, have a mean of 0, and are uncorrelated with the person saliences, v_{ij} and w_{ik}), for a particular time of measurement, the variance of $x_i(s)$ is given by

$$\sigma_x^2(s) = \sum_{j=1}^{r_1} \sum_{j'=1}^{r_1} g_j(s) g_{j'}(s) \sigma_{v_j v_{j'}} + \sigma_E^2(s), \quad (6)$$

and the variance of $y_i(t)$ is given by

$$\sigma_Y^2(t) = \sum_{k=1}^{r_0} \sum_{k'=1}^{r_0} h_k(t) h_{k'}(t) \sigma_{w_k w_{k'}} + \sigma_F^2(t), \quad (7)$$

where $\sigma_E^2(s)$ and $\sigma_F^2(t)$ are the error variances for $e_i(s)$ and $f_i(t)$ at times s and t , respectively, and where $\sigma_{v_j v_{j'}}$ and $\sigma_{w_k w_{k'}}$ are the covariances of the person weights, v_{ij} with $v_{ij'}$ and w_{ik} with $w_{ik'}$, respectively.

Similarly, for particular times s and t , the covariance of X and Y is given by

$$\sigma_{X(s)Y(t)} = \sum_{j=1}^{r_1} \sum_{k=1}^{r_0} g_j(s) h_k(t) \sigma_{v_j w_k}, \quad (8)$$

where $\sigma_{X(s)Y(t)}$ is the covariance between the observed variables, $X_i(s)$ and $Y_i(t)$, and $\sigma_{v_j w_k}$ is the covariance between the person weights, v_{ij} and w_{ik} .

After the removal of $\sigma_E^2(s)$ and $\sigma_F^2(t)$, Equations 6 and 7 may be substituted into Equation 1; this substitution will lead to a dynamic model of reliability,

$$\rho_R[X(s)] \equiv \rho_X(s) = \frac{\sum_{j=1}^{r_1} \sum_{j'=1}^{r_1} g_j(s) g_{j'}(s) \sigma_{v_j v_{j'}}}{\sum_{j=1}^{r_1} \sum_{j'=1}^{r_1} g_j(s) g_{j'}(s) \sigma_{v_j v_{j'}} + \sigma_E^2(s)} \quad (9)$$

and

$$\rho_R[Y(t)] \equiv \rho_Y(t) = \frac{\sum_{k=1}^{r_0} \sum_{k'=1}^{r_0} h_k(t) h_{k'}(t) \sigma_{w_k w_{k'}}}{\sum_{k=1}^{r_0} \sum_{k'=1}^{r_0} h_k(t) h_{k'}(t) \sigma_{w_k w_{k'}} + \sigma_F^2(t)}. \quad (10)$$

Note that the longitudinal model in Equations 9 and 10 permits the assessment of the reliability as a single index at each point of measurement; that is, the reliability of $X(s)$ has the set of values $\{\rho_X(s_1), \rho_X(s_2), \dots, \rho_X(s_p)\}$ for the p th measurement period, and the reliability of $Y(t)$ has the set of values $\{\rho_Y(t_1), \rho_Y(t_2), \dots, \rho_Y(t_q)\}$ for the q th measurement period.

By substituting Equations 6, 7, and 8 into Equation 2, a dynamic model of validity is given by

$$\rho_v(s, t) \equiv |\rho_{x(s)y(t)}| = \frac{\left| \sum_{j=1}^{r_1} \sum_{k=1}^{r_0} g_j(s) h_k(t) \sigma_{v_j w_k} \right|}{\left[\sum_{j=1}^{r_1} \sum_{j'=1}^{r_1} g_j(s) g_{j'}(s) \sigma_{v_j v_{j'}} + \sigma_E^2(s) \right]^{1/2} \left[\sum_{k=1}^{r_0} \sum_{k'=1}^{r_0} h_k(t) h_{k'}(t) \sigma_{w_k w_{k'}} + \sigma_F^2(t) \right]^{1/2}} \quad (11)$$

Note that the longitudinal model presented in Equation 11 permits the assessment of the validities for all pairs of times for the two variables. When the periods of measurement coincide, the coefficient represents concurrent validity; when they differ, predictive validities are obtained.

Additionally, note that constancy and change are reflected in the latent basis curves, $\{g_j(s)\}$ and $\{h_k(t)\}$ from X and Y , respectively. For example, if in the previous illustration, $r_1 = 1$, $g_1(s) \equiv 1$, and $v_1(s) = a_i$, then Equation 3 would become $x_i(s) = a_i + e_i(s)$. This equation represents the case in which there are individual differences as reflected in a_i , but there is no change (i.e., the attribute remains constant across the time periods assessed).

Matrix Formulation

The above scalar equations have the following matrix representation. Let the random vectors of observed variables, \mathbf{x} and \mathbf{y} , be defined as

$$\mathbf{x}' = [X(s_1), X(s_2), \dots, X(s_p)] \quad (12)$$

and

$$\mathbf{y}' = [Y(t_1), Y(t_2), \dots, Y(t_q)] \quad (13)$$

The mean vector and covariance matrix of the supermatrix formed by the concatenation of \mathbf{y} and \mathbf{x} are given by

$$E \begin{bmatrix} \mathbf{y} \\ \mathbf{x} \end{bmatrix} = \begin{bmatrix} \boldsymbol{\mu}_y \\ \boldsymbol{\mu}_x \end{bmatrix} \quad (14)$$

and

$$\text{COV} \begin{bmatrix} \mathbf{y} \\ \mathbf{x} \end{bmatrix} = \begin{bmatrix} \boldsymbol{\Sigma}_{yy'} & \boldsymbol{\Sigma}_{yx'} \\ \boldsymbol{\Sigma}_{xy'} & \boldsymbol{\Sigma}_{xx'} \end{bmatrix}, \quad (15)$$

where $E[\cdot]$ and $\text{COV}[\cdot]$ are the expectation and covariance operators, respectively.

Further, define the random vectors of person weights, \mathbf{v} and \mathbf{w} , as

$$\mathbf{v}' = [V_1, V_2, \dots, V_{r_1}] \quad (16)$$

and

$$\mathbf{w}' = [W_1, W_2, \dots, W_{r_0}]. \quad (17)$$

Also, define the random error vectors, \mathbf{e} and \mathbf{f} , as

$$\mathbf{e}' = [E(s_1), E(s_2), \dots, E(s_p)] \quad (18)$$

and

$$\mathbf{f}' = [F(t_1), F(t_2), \dots, F(t_q)]. \quad (19)$$

Let Γ_0 be a $(q \times r_0)$ matrix and Γ_1 be a $(p \times r_1)$ matrix containing the elements $\{h_k(t)\}$ and $\{g_j(s)\}$, respec-

tively. The basic modeling Equations (Equations 3, 4, 6, 7, and 8) are then rewritten as

$$\begin{bmatrix} \mathbf{y} \\ \mathbf{x} \end{bmatrix} = \begin{bmatrix} \Gamma_0 & 0 \\ 0 & \Gamma_1 \end{bmatrix} \begin{bmatrix} \mathbf{w} \\ \mathbf{v} \end{bmatrix} + \begin{bmatrix} \mathbf{f} \\ \mathbf{e} \end{bmatrix}. \quad (20)$$

With the standard assumptions of errors of measurement, the observed or measured means can be written as

$$\begin{bmatrix} \mu_y \\ \mu_x \end{bmatrix} = \begin{bmatrix} \Gamma_0 & 0 \\ 0 & \Gamma_1 \end{bmatrix} \begin{bmatrix} \mu_w \\ \mu_v \end{bmatrix} \quad (21)$$

and the observed covariances as

$$\begin{bmatrix} \Sigma_{yy'} & \Sigma_{yx'} \\ \Sigma_{xy'} & \Sigma_{xx'} \end{bmatrix} = \begin{bmatrix} \Gamma_0 & 0 \\ 0 & \Gamma_1 \end{bmatrix} \begin{bmatrix} \Sigma_{ww'} & \Sigma_{wv'} \\ \Sigma_{vw'} & \Sigma_{vv'} \end{bmatrix} \begin{bmatrix} \Gamma_0' & 0 \\ 0 & \Gamma_1' \end{bmatrix} + \begin{bmatrix} \mathbf{D}_{ff'} & 0 \\ 0 & \mathbf{D}_{ee'} \end{bmatrix}, \quad (22)$$

where

μ_w and μ_v are the mean vectors of the person weights, w and v , respectively;

$\Sigma_{ww'}$, $\Sigma_{vv'}$, and $\Sigma_{wv'}$ are the covariance matrices of the saliences, w with w' , v with v' , and v with w' , respectively; and

$\mathbf{D}_{ff'}$ and $\mathbf{D}_{ee'}$ are diagonal matrices of measurement error variances.

In this form, Equations 20, 21, and 22 are the basic equations of a confirmatory factor analytic model (Bock & Bargmann, 1966; Jöreskog, 1969).

Estimation and Hypothesis Testing

Let the estimates of the population mean vectors (μ_x and μ_y) and the population covariance matrices ($\Sigma_{xx'}$, $\Sigma_{yy'}$, and $\Sigma_{xy'}$) be obtained by the usual sample statistics of the mean vectors, $\hat{\mu}_x$ and $\hat{\mu}_y$, and by the usual unbiased sample statistics of the covariance matrices, $\hat{\Sigma}_{xx'}$, $\hat{\Sigma}_{yy'}$, and $\hat{\Sigma}_{xy'}$, respectively (Morrison, 1990, pp. 98 and 100, equations 3 and 9). Using these estimated means and covariances, and with proper identification constraints, all of the parameters of the model (μ_w , μ_v , $\Sigma_{ww'}$, $\Sigma_{vv'}$, $\Sigma_{wv'}$, Γ_0 , Γ_1 , $\mathbf{D}_{ee'}$, and $\mathbf{D}_{ff'}$) can be estimated by any computer program that performs confirmatory factor analysis, such as LISREL 7 (Jöreskog & Sörbom, 1989), CALIS (SAS Institute, Inc., 1989), or EQS (Bentler, 1989). Additionally, if the multivariate normality assumption is tenable, maximum likelihood estimates and asymptotic χ^2 tests can be obtained. With these parameter estimates, the following vectors of reliability coefficients are obtained

$$\rho_x = \mathbf{D}[\Gamma_1 \Sigma_{vv'} \Gamma_1'] \mathbf{D}^{-1} [\Gamma_1 \Sigma_{vv'} \Gamma_1' + \mathbf{D}_{ee'}] \mathbf{1}_p \quad (23)$$

and

$$\rho_y = \mathbf{D}[\Gamma_0 \Sigma_{ww'} \Gamma_0'] \mathbf{D}^{-1} [\Gamma_0 \Sigma_{ww'} \Gamma_0' + \mathbf{D}_{ff'}] \mathbf{1}_q, \quad (24)$$

where $\mathbf{D}[\cdot]$ is a diagonal matrix containing the principal diagonal elements of the matrix $[\cdot]$. Note that ρ_x and ρ_y are vectors of order p and q that contain the variable's reliability at each time point.

A $(p \times q)$ matrix of validity coefficients may be determined by

$$\mathbf{P}_{xy} = \mathbf{D}^{-1/2} [\Gamma_1 \Sigma_{vv'} \Gamma_1' + \mathbf{D}_{ee'}]^{-1} |\Gamma_1 \Sigma_{vv'} \Gamma_1'| \mathbf{D}^{-1/2} [\Gamma_0 \Sigma_{ww'} \Gamma_0' + \mathbf{D}_{ff'}], \quad (25)$$

where $|\Gamma_1 \Sigma_{vv'} \Gamma_1'|$ is a matrix that contains the absolute values of the elements of the matrix $\Gamma_0 \Sigma_{vv'} \Gamma_1'$. Elements of \mathbf{P}_{xy} are specific validities of X at time s and Y at time t [i.e., $\rho_{y(s,t)}$]. Further, these coefficients may be disattenuated (if desired) by

$$\mathbf{P}_{xy} = \mathbf{D}^{-1/2} [\Gamma_1 \Sigma_{vv'} \Gamma_1'] [\Gamma_1 \Sigma_{vv'} \Gamma_1']^{-1} \mathbf{D}^{-1/2} [\Gamma_0 \Sigma_{ww'} \Gamma_0']. \quad (26)$$

Note that \mathbf{P}_{xy} can contain both concurrent and predictive validities as elements and that \mathbf{P}_{xy} will contain elements that can be considered construct (concurrent and predictive) validities.

Extensions

Congeneric Analog

Jöreskog (1971) proposed a congeneric measurement model that included parallel, τ -equivalent, and essentially τ -equivalent models as special cases. Recently, using mean structure, Millsap & Everson (1991) refined a number of submodels. In the notation used here, this congeneric or linearly equivalent model is given by

$$x_i(m) = \alpha(m) + g(m)v_i + e_i(m), \quad (27)$$

with

$$\mu_{x(m)} = \alpha(m) + g(m)\mu_v, \quad (28)$$

$$\sigma_{x(m)}^2 = g(m)^2 \sigma_v^2 + \sigma_{E(m)}^2, \quad (29)$$

and

$$\sigma_{x(m)x(m')} = g(m)g(m')\sigma_v^2, \quad (30)$$

where $m \in \{1, 2, \dots, p\}$ and in this situation it indexes variables instead of time, and where $\alpha(m)$ is an additive constant. Because of identifiability constraints, this linearly equivalent representation may appear differently (although equivalently) in other contexts. Specifically, the above model would be identified by a marker variable; that is, for a particular variable, m^* , $\alpha(m^*) = 0$, and $g(m^*) = 1$.

Meredith & Tisak (1990) proposed a reparameterization of this congeneric model for developmental applications:

$$x_i(s) = \alpha(s) + g(s)v_i + e_i(s) \quad (31)$$

and

$$E[X(s)] = \alpha(s) + g(s)\mu_v, \quad (32)$$

where s indexes time and $\alpha(s)$ is an additive period/practice effect [with $\alpha(1) = 0$ for identification]. Although Equations 27 and 28 are identical in form to Equations 31 and 32, they are very different in function. Specifically, Equations 27 and 28 depict a number of similar measures at one time; Equations 31 and 32 represent the repeated measurement of a single variable.

Generalizations to Multiple Populations

Although validity was not considered, the extension to multiple populations was proposed by Meredith & Tisak (1990) and the extension to multivariate situations was presented by Tisak & Meredith (1990). The incorporation of the ideas given here should be straightforward. An example is provided below for this case.

Test-Retest Reliability

When there are only two time periods, the correlation between the measurements obtained at these times has been used as an estimate of reliability (Lord & Novick, 1968). This procedure has been criticized (Heise, 1969) from a sociological perspective. From the approach presented here, it may be asked what the latent curve models of Equation 3 are such that the test-retest correlation has the form depicted by the reliability coefficient in Equation 1 (i.e., the test-retest correlation meets the definition of the reliability coefficient). For two time points, there are only two models that meet the above criterion.

The first model is the longitudinal version as an essentially τ -equivalent measurement model with equal error variances (Lord & Novick, 1968). The mean and covariance structures are given by

$$\begin{bmatrix} \mu_{x_1} \\ \mu_{x_2} \end{bmatrix} = \begin{bmatrix} 0 \\ \alpha(2) \end{bmatrix} + \begin{bmatrix} 1 \\ 1 \end{bmatrix} \mu_v \quad (33)$$

and

$$\begin{bmatrix} \sigma_{x_1}^2 & \sigma_{x_1 x_2} \\ \sigma_{x_2 x_1} & \sigma_{x_2}^2 \end{bmatrix} = \begin{bmatrix} 1 \\ 1 \end{bmatrix} \sigma_v^2 [1 \ 1] + \begin{bmatrix} \sigma_E^2 & 0 \\ 0 & \sigma_E^2 \end{bmatrix}, \quad (34)$$

respectively. The graph of the regression of the predicted X on time consists of a series of parallel lines whose constant slope is given by an additive increase at Time 2, $\alpha(2)$. Psychologically, this model represents persons who have different initial values on an attribute, but then change at identical rates.

The second model, a submodel of the former, is the longitudinal version of a parallel measurement model (Lord & Novick, 1968) and has the following mean and covariance structure:

$$\begin{bmatrix} \mu_{x_1} \\ \mu_{x_2} \end{bmatrix} = \begin{bmatrix} 1 \\ 1 \end{bmatrix} \mu_v \quad (35)$$

and

$$\begin{bmatrix} \sigma_{x_1}^2 & \sigma_{x_1 x_2} \\ \sigma_{x_2 x_1} & \sigma_{x_2}^2 \end{bmatrix} = \begin{bmatrix} 1 \\ 1 \end{bmatrix} \sigma_v^2 [1 \ 1] + \begin{bmatrix} \sigma_E^2 & 0 \\ 0 & \sigma_E^2 \end{bmatrix}, \quad (36)$$

respectively. The graph of the regression of the predicted X on time consists of a series of lines that are parallel and horizontal from Time 1 to Time 2. This submodel depicts persons who differ in initial status but subsequently do not change. Thus, from a developmental framework, the test-retest correlation can represent a reliability coefficient only if there are equal error variances and, at most, an additive change effect. In particular, only a very restrictive form of change is possible; that is, persons may differ in initial status, but they must change at the same rate (which includes no change).

Examples

Sociological Example

This example used longitudinal sociological variables collected at three time points (1966, 1967, 1971). These data ($N = 932$) were reported by Wheaton et al. (1977). A more complete description is given by Summers, Hough, Scott, & Folse (1969). The respondent's occupational status, as measured by the Duncan Socioeconomic Index (SEI), and their educational attainment were the variables of interest. Note that SEI is longitudinal, and education is static. The correlations, standard deviations (SDs), and means for these variables are given in Table 1.

The longitudinal model selected to represent these data was

$$x_i(s) = v_i \quad (37)$$

Table 1
 Correlations, Standard Deviations, and Means
 for the Sociological Dataset

Dataset and Statistic	Y_1	Y_2	Y_3	X
Y_1 : SEI (1966)	1.00			
Y_2 : SEI (1967)	.82	1.00		
Y_3 : SEI (1971)	.81	.78	1.00	
X : Education (1966)	.54	.53	.54	1.00
Mean	37.49	36.72	37.47	10.90
Standard Deviation	21.22	21.00	20.98	3.10

and

$$y_i(t) = h(t)w_i + f_i(t), \quad (38)$$

with error variance $\sigma_f^2(t)$,

where

$$i \in \{1, 2, \dots, 932\},$$

$$s = 1, \text{ and}$$

$$t \in \{1, 2, 3\}.$$

Because education was measured only in 1966, the model for $x_i(t)$ was predetermined. However, for the SEI a number of models for $y_i(t)$ were possible. The above model was selected as best by the χ^2 statistic; the χ^2 with 2 degrees of freedom (df) was 4.83, $p = .089$. Other possible approaches, such as $\sigma_f^2(t) = \sigma_f^2$ or $h(t) \equiv 1$, proved to be statistically significant. All analyses were performed by LISREL 7 (Jöreskog & Sörbom, 1989).

After preliminary analyses determined the best developmental model for these data, further analyses were performed on the joint set of variables, $\{Y_1, Y_2, Y_3, X\}$. In matrix notation, the following estimates were obtained:

$$\begin{bmatrix} \hat{\Gamma}_0 & \mathbf{0} \\ \mathbf{0} & \hat{\Gamma}_1 \end{bmatrix} = \begin{bmatrix} 1.000 & 0.000 \\ .976 & 0.000 \\ .990 & 0.000 \\ 0.000 & 1.000 \end{bmatrix}, \quad (39)$$

$$\begin{bmatrix} \hat{\mu}_w \\ \hat{\mu}_v \end{bmatrix} = \begin{bmatrix} 37.640 \\ 10.900 \end{bmatrix}, \quad (40)$$

$$\begin{bmatrix} \hat{\Sigma}_{ww'} & \hat{\Sigma}_{wv'} \\ \hat{\Sigma}_{vw'} & \hat{\Sigma}_{vv'} \end{bmatrix} = \begin{bmatrix} 366.029 & 35.422 \\ 35.422 & 9.600 \end{bmatrix}, \quad (41)$$

and

$$\begin{bmatrix} \hat{D}_r & \mathbf{0} \\ \mathbf{0}' & \hat{D}_{ee'} \end{bmatrix} = \begin{bmatrix} 72.505 & 0.000 & 0.000 & 0.000 \\ 0.000 & 91.937 & 0.000 & 0.000 \\ 0.000 & 0.000 & 96.092 & 0.000 \\ 0.000 & 0.000 & 0.000 & 0.000 \end{bmatrix}. \quad (42)$$

The goodness of fit for this model was $\chi^2(4 df) = 6.39, p = .172$; thus, the data were represented quite well.

Reliability for the SEI was obtained as

$$\hat{\rho}_y = \mathbf{D}[\hat{\Gamma}_0 \hat{\Sigma}_{ww'} \hat{\Gamma}_0'] \mathbf{D}^{-1}[\hat{\Gamma}_0 \hat{\Sigma}_{ww'} \hat{\Gamma}_0' + \hat{D}_r] \mathbf{1}_3 = \begin{bmatrix} .835 \\ .791 \\ .789 \end{bmatrix}. \quad (43)$$

The vector of validity coefficients was

$$\hat{\rho}_{xy} = \mathbf{D}^{-1/2}[\hat{\Gamma}_1 \hat{\Sigma}_{vv'} \hat{\Gamma}_1'] \hat{\Gamma}_1 \hat{\Sigma}_{ww'} \hat{\Gamma}_0' \mathbf{D}^{-1/2}[\hat{\Gamma}_0 \hat{\Sigma}_{ww'} \hat{\Gamma}_0' + \hat{D}_r] = [.546 \quad .532 \quad .531]. \quad (44)$$

Finally, these validity coefficients were disattenuated (for measurement error in Y), and the results were

$$\hat{\rho}_{xty} = \mathbf{D}^{-1/2}[\hat{\Gamma}_1 \hat{\Sigma}_{vv'} \hat{\Gamma}_1'] \hat{\Gamma}_1 \hat{\Sigma}_{ww'} \hat{\Gamma}_0' \mathbf{D}^{-1/2}[\hat{\Gamma}_0 \hat{\Sigma}_{ww'} \hat{\Gamma}_0'] = [.598 \quad .598 \quad .598]. \quad (45)$$

Note that the disattenuated validity coefficients were all equal. In general, this will always be true when the variables are highly stable (i.e., when $r_0 = r_1 = 1$).

Psychological Example

For the psychological example, a dataset with some unique and interesting features was selected. The original study used a cohort sequential design (Schaie, 1965). Thus, by using a cohort sequential design, the data were longitudinal in multiple populations. Also, some of the age groups overlapped across groups. Thus, it was possible to lag the latent basis curves, which in turn permitted these curves to be invariant across the populations (Meredith & Tisak, 1990).

Specifically, the data used here were part of the data collected by Nesselroade & Baltes (1974) in 1970, 1971, and 1972 on the number-correct scores of the Number Facility (NUMF) and Number Series (NUMS) subtests of the Primary Mental Abilities Test (Thurstone & Thurstone, 1962). In the present context, NUMS could be evaluated as a proxy for NUMF (an arithmetic test).

For the purpose of this analysis, the focus was on only two of the male birth-year cohorts. Cohort 1 ($N = 98$) was assessed at grades 7, 8, and 9. Cohort 2 ($N = 93$) was measured at grades 8, 9, and 10. Note that the ages were fairly homogeneous within grade. The correlations, SDs, and means for these variables and cohorts are given in Table 2.

Table 2
 Correlations, Standard Deviations (SDs), and Means for Cohort 1 (Lower Triangle) and Cohort 2 (Upper Triangle) of the Psychological Dataset

Dataset and Statistic	Y_1	Y_2	Y_3	X_1	X_2	X_3	Mean	SD
Y_1 : NUMF (1970)	—	.77	.73	.55	.59	.63	4.51	11.33
Y_2 : NUMF (1971)	.63	—	.75	.54	.63	.69	5.36	13.86
Y_3 : NUMF (1972)	.61	.71	—	.55	.66	.75	6.15	15.02
X_1 : NUMS (1970)	.46	.53	.45	—	.59	.57	2.91	6.19
X_2 : NUMS (1971)	.45	.55	.48	.53	—	.74	3.22	7.42
X_3 : NUMS (1972)	.61	.59	.60	.55	.66	—	3.53	8.22
Mean	4.19	5.00	5.02	2.37	3.03	3.21		
SD	10.19	12.70	13.89	5.61	6.70	7.96		

The longitudinal model selected to represent these data was

$$x_i^{(k)}(s) = g^{(k)}(s)v_i^{(k)} + e_i^{(k)}(s), \tag{46}$$

with error variance $\sigma_E^{2(k)}(s)$, and

$$y_i^{(k)}(t) = h^{(k)}(t)w_i^{(k)} + f_i^{(k)}(t), \tag{47}$$

with error variance $\sigma_F^{2(k)}(t)$,

where

$k \in \{1, 2\}$ represents Cohort 1 or 2,

s and $t \in \{1, 2, 3\}$ represent year evaluated (1970, 1971, and 1972), and

$i \in \{1, 2, \dots, 98\}$, if $k = 1$, or

$i \in \{1, 2, \dots, 93\}$, if $k = 2$.

The single latent basis curve was lagged or constrained; that is, $g^{(1)}(2) = g^{(2)}(1)$, $g^{(1)}(3) = g^{(2)}(2)$ for NUMS, and $h^{(1)}(2) = h^{(2)}(1)$, $h^{(1)}(3) = h^{(2)}(2)$ for NUMF.

After preliminary analyses determined the best developmental model for these data, further analyses were performed on the joint set of variables, $\{Y_1, Y_2, Y_3, X_1, X_2, X_3\}$. Again, all analyses were performed by

LISREL7 (Jöreskog & Sörbom, 1989). In matrix notation, the following estimates were obtained:

$$\begin{bmatrix} \hat{\Gamma}_0^{(1)} & \mathbf{0} \\ \mathbf{0} & \hat{\Gamma}_1^{(1)} \end{bmatrix} = \begin{bmatrix} 1.000 & 0.000 \\ 1.209 & 0.000 \\ 1.397 & 0.000 \\ 0.000 & 1.000 \\ 0.000 & 1.205 \\ 0.000 & 1.444 \end{bmatrix} \quad (48)$$

and

$$\begin{bmatrix} \hat{\Gamma}_0^{(2)} & \mathbf{0} \\ \mathbf{0} & \hat{\Gamma}_1^{(2)} \end{bmatrix} = \begin{bmatrix} 1.209 & 0.000 \\ 1.397 & 0.000 \\ 1.563 & 0.000 \\ 0.000 & 1.205 \\ 0.000 & 1.444 \\ 0.000 & 1.613 \end{bmatrix} \quad (49)$$

From the above matrices, note that the latent basis curve for NUMF, $g(s) \in \{1.000, 1.209, 1.397, 1.563\}$, increased as grade in school changed $\{7, 8, 9, 10\}$. Additionally, the latent basis curve for NUMS, $h(t) \in \{1.000, 1.205, 1.444, 1.613\}$, also increased as grade in school changed $\{7, 8, 9, 10\}$. Thus, for both NUMF and NUMS, positive change or learning occurred across the four grades.

For each population or cohort, the means and the covariances of the saliences, v and w , were

$$\begin{bmatrix} \hat{\mu}_w^{(1)} \\ \hat{\mu}_v^{(1)} \end{bmatrix} = \begin{bmatrix} 10.187 \\ 5.548 \end{bmatrix}, \quad (50)$$

$$\begin{bmatrix} \hat{\mu}_w^{(2)} \\ \hat{\mu}_v^{(2)} \end{bmatrix} = \begin{bmatrix} 9.648 \\ 5.119 \end{bmatrix}, \quad (51)$$

$$\begin{bmatrix} \hat{\Sigma}_{ww}^{(1)} & \hat{\Sigma}_{wv}^{(1)} \\ \hat{\Sigma}_{vw}^{(1)} & \hat{\Sigma}_{vv}^{(1)} \end{bmatrix} = \begin{bmatrix} 10.122 & 4.945 \\ 4.945 & 3.323 \end{bmatrix}, \quad (52)$$

and

$$\begin{bmatrix} \hat{\Sigma}_{ww}^{(2)} & \hat{\Sigma}_{wv}^{(2)} \\ \hat{\Sigma}_{vw}^{(2)} & \hat{\Sigma}_{vv}^{(2)} \end{bmatrix} = \begin{bmatrix} 11.046 & 5.521 \\ 5.521 & 3.417 \end{bmatrix}. \quad (53)$$

Finally, the error variances for each cohort were

$$\begin{bmatrix} \hat{D}_{rr}^{(1)} & \mathbf{0} \\ \mathbf{0} & \hat{D}_{ee}^{(1)} \end{bmatrix} = \begin{bmatrix} 7.489 & 0.000 & 0.000 & 0.000 & 0.000 & 0.000 \\ 0.000 & 7.788 & 0.000 & 0.000 & 0.000 & 0.000 \\ 0.000 & 0.000 & 7.834 & 0.000 & 0.000 & 0.000 \\ 0.000 & 0.000 & 0.000 & 2.978 & 0.000 & 0.000 \\ 0.000 & 0.000 & 0.000 & 0.000 & 3.993 & 0.000 \\ 0.000 & 0.000 & 0.000 & 0.000 & 0.000 & 2.964 \end{bmatrix} \quad (54)$$

and

$$\begin{bmatrix} \hat{\mathbf{D}}_{rr}^{(2)} & \mathbf{0} \\ \mathbf{0} & \hat{\mathbf{D}}_{ee}^{(2)} \end{bmatrix} = \begin{bmatrix} 5.657 & 0.000 & 0.000 & 0.000 & 0.000 & 0.000 \\ 0.000 & 6.692 & 0.000 & 0.000 & 0.000 & 0.000 \\ 0.000 & 0.000 & 9.166 & 0.000 & 0.000 & 0.000 \\ 0.000 & 0.000 & 0.000 & 4.441 & 0.000 & 0.000 \\ 0.000 & 0.000 & 0.000 & 0.000 & 3.197 & 0.000 \\ 0.000 & 0.000 & 0.000 & 0.000 & 0.000 & 2.988 \end{bmatrix}. \quad (55)$$

The goodness of fit for this model was $\chi^2(26) = 30.82, p = .235$; thus, the data were represented quite well.

Reliability for NUMF was obtained for each cohort by

$$\hat{\rho}_y^{(1)} = \mathbf{D}[\hat{\Gamma}_0 \hat{\Sigma}_{ww}^{(1)} \hat{\Gamma}'_0] \mathbf{D}^{-1}[\hat{\Gamma}_0 \hat{\Sigma}_{ww}^{(1)} \hat{\Gamma}'_0 + \hat{\mathbf{D}}_{rr}^{(1)}] \mathbf{1}_3 = \begin{bmatrix} .575 \\ .655 \\ .716 \end{bmatrix} \quad (56)$$

and

$$\hat{\rho}_y^{(2)} = \mathbf{D}[\hat{\Gamma}_0 \hat{\Sigma}_{ww}^{(2)} \hat{\Gamma}'_0] \mathbf{D}^{-1}[\hat{\Gamma}_0 \hat{\Sigma}_{ww}^{(2)} \hat{\Gamma}'_0 + \hat{\mathbf{D}}_{rr}^{(2)}] \mathbf{1}_3 = \begin{bmatrix} .740 \\ .763 \\ .746 \end{bmatrix}. \quad (57)$$

Reliability for NUMS was obtained for each cohort by

$$\hat{\rho}_x^{(1)} = \mathbf{D}[\hat{\Gamma}_1 \hat{\Sigma}_{vv}^{(1)} \hat{\Gamma}'_1] \mathbf{D}^{-1}[\hat{\Gamma}_1 \hat{\Sigma}_{vv}^{(1)} \hat{\Gamma}'_1 + \hat{\mathbf{D}}_{ee}^{(1)}] \mathbf{1}_3 = \begin{bmatrix} .527 \\ .547 \\ .700 \end{bmatrix} \quad (58)$$

and

$$\hat{\rho}_x^{(2)} = \mathbf{D}[\hat{\Gamma}_1 \hat{\Sigma}_{vv}^{(2)} \hat{\Gamma}'_1] \mathbf{D}^{-1}[\hat{\Gamma}_1 \hat{\Sigma}_{vv}^{(2)} \hat{\Gamma}'_1 + \hat{\mathbf{D}}_{ee}^{(2)}] \mathbf{1}_3 = \begin{bmatrix} .527 \\ .547 \\ .700 \end{bmatrix} \quad (59)$$

The matrices of validity coefficients were

$$\hat{\mathbf{P}}_{xy}^{(1)} = \mathbf{D}^{-1/2}[\hat{\Gamma}_1 \hat{\Sigma}_{vv}^{(1)} \hat{\Gamma}'_1 + \hat{\mathbf{D}}_{ee}^{(1)}] \hat{\Gamma}_1 \hat{\Sigma}_{ww}^{(1)} \hat{\Gamma}'_0 \mathbf{D}^{-1/2}[\hat{\Gamma}_0 \hat{\Sigma}_{ww}^{(1)} \hat{\Gamma}'_0 + \hat{\mathbf{D}}_{rr}^{(1)}] = \begin{bmatrix} .469 & .501 & .524 \\ .478 & .511 & .534 \\ .541 & .578 & .604 \end{bmatrix} \quad (60)$$

and

$$\hat{\mathbf{P}}_{xy}^{(2)} = \mathbf{D}^{-1/2}[\hat{\Gamma}_1 \hat{\Sigma}_{vv}^{(2)} \hat{\Gamma}'_1 + \hat{\mathbf{D}}_{ee}^{(2)}] \hat{\Gamma}_1 \hat{\Sigma}_{ww}^{(2)} \hat{\Gamma}'_0 \mathbf{D}^{-1/2}[\hat{\Gamma}_0 \hat{\Sigma}_{ww}^{(2)} \hat{\Gamma}'_0 + \hat{\mathbf{D}}_{rr}^{(2)}] = \begin{bmatrix} .562 & .570 & .564 \\ .642 & .652 & .645 \\ .669 & .679 & .672 \end{bmatrix}, \quad (61)$$

for Cohorts 1 and 2, respectively.

Finally, these validity coefficients were disattenuated (for measurement error in Y and X), and the results were

$$\hat{\mathbf{P}}_{xy}^{(1)} = \mathbf{D}^{-1/2}[\hat{\Gamma}_1 \hat{\Sigma}_{vv}^{(1)} \hat{\Gamma}'_1] \hat{\Gamma}_1 \hat{\Sigma}_{ww}^{(1)} \hat{\Gamma}'_0 \mathbf{D}^{-1/2}[\hat{\Gamma}_0 \hat{\Sigma}_{ww}^{(1)} \hat{\Gamma}'_0] = .853 \mathbf{U}_3, \quad (62)$$

and

$$\hat{P}_{x,y}^{(2)} = \mathbf{D}^{-1/2} [\hat{\Gamma}_1 \hat{\Sigma}_{vw}^{(2)} \hat{\Gamma}_1'] [\hat{\Gamma}_1 \hat{\Sigma}_{vw}^{(2)} \hat{\Gamma}_0' | \mathbf{D}^{-1/2} [\hat{\Gamma}_0 \hat{\Sigma}_{vw}^{(2)} \hat{\Gamma}_0']] = .899 \mathbf{U}_3, \quad (63)$$

for Cohorts 1 and 2, respectively, where \mathbf{U}_3 is the unit matrix of order 3 (i.e., $\mathbf{U}_3 = \mathbf{1}_3 \mathbf{1}'_3$). Note that the disattenuated validity coefficients were all equal. In general, this will always be true when the variables are highly stable (i.e., when $r_0 = r_1 = 1$).

Discussion

The LCA approach developed here incorporates a number of models. Specifically, LCA is similar to Jöreskog's (1971) congeneric measurement models if different measurement periods are substituted for different variables. Additionally, practice effects may be represented as an additive constant. Finally, the problem extends naturally to include multiple populations or groups.

Although it was not the intent of this paper to compare the proposed explication of longitudinal reliability and validity to other approaches [e.g., Wheaton et al. (1977)], one formulation (Willett, 1989) is similar enough to warrant a brief comment. Willett suggests a "...population reliability of the estimated growth-rate..." (p. 592). Willett's basic model is $x_i(s) = a_i + b_i s + e_i(s)$; that is, a linear model that includes the individual difference parameters, initial status (a_i) and rate (b_i). Further, he attempts to decompose the observed rate variability into interindividual rate variability and intraindividual sampling rate variability, and to form an index of interindividual variance to observed rate variance. Unfortunately, there are confounds, which include measurement error, in the model. Hence, there are some notable differences between the two approaches: (1) classical definitions of reliability and validity are nested within the proposed framework, which is not apparent in Willett's work; (2) Willett focused his definition on the growth-rate variable, but here the instrument or measure is the focus; and (3) the approach suggested here is more general in scope, both in estimation and testing (i.e., using structural equations modeling) and in the incorporation of nonlinear longitudinal functions.

Following a similar direction, Collins (1991) and others (Cattell, 1964; Collins & Cliff, 1990; Cronbach, Gleser, Nanda, & Rajaratnam, 1972; Rogosa, Brandt, & Zimowski, 1982) have criticized reliability as defined by classical test theory for being too focused on interindividual differences and ignoring intraindividual differences. Although a full exposition on this topic is beyond the scope of this paper, it is useful to consider how the proposed technique is related to the intraindividual-interindividual differences discussion.

First, the depiction of dynamic reliability in Equations 9 and 10 is consistent with the classical approach in Equation 1. Clearly, both are based on interindividual differences. What would happen if everyone changed identically, that is, if there were no individual differences? To explore this situation, extract the vectors and matrices associated with the vector \mathbf{x} from Equations 20–22 and impose that $\mathbf{v} = \mu_v$ [i.e., everyone has the same saliences, which implies that the covariance matrix of \mathbf{v} is $\mathbf{0}$ ($\Sigma_{vv} = \mathbf{0}$)], which yields

$$\mathbf{x} = \Gamma_1 \mu_v + \mathbf{e}, \quad (64)$$

$$\mu_x = \Gamma_1 \mu_v, \quad (65)$$

and

$$\Sigma_{xx'} = \mathbf{D}_{ee'}. \quad (66)$$

Substituting Equations 64 and 65 into Equation 23 confirms that the reliability is 0 at each time of measurement (i.e., the vector $\rho_x = \mathbf{0}$). However, also observe that the latent trajectory for each and all persons is given by $\mu_x = \Gamma_1 \mu_v$. Further, the standard error of measurement is given by the square root of the variance of the errors, $\mathbf{D}_{ee'}^{1/2}$. Hence, if interest is in the precision of the latent trajectory in this atypical situation, $\mathbf{D}_{ee'}^{1/2}$ could be used, or equivalently, $\sigma_E(s)$ for times $s \in \{s_1, s_2, \dots, s_p\}$ to construct (confidence) bands around the trajectory.

Thus, although the dynamic reliability approach embraces the idea that reliability is defined across persons, it not only permits, but encourages, the statistical testing of a great variety of submodels that includes models of change and constancy without individual differences. Hence, researchers who use this technique must consider and understand the implications of both the reliability coefficients and the longitudinal model: Because the reliability coefficient is a function of both measurement error and individual differences, the appropriate longitudinal model must first be determined before interpreting this coefficient.

Finally, two theoretical issues have been addressed. The first question is what type of development or change is required for the test-retest correlation to serve as a surrogate for the reliability coefficient. For two measurement periods, it has been suggested that the two are reasonably equivalent if a longitudinal version of either an essentially τ -equivalent model with equal error variances or a parallel model is present. Practically, this means that all persons must change at the same rate, but they may differ in initial status. The second issue is what type of change will make (disattenuated) concurrent and predictive validities identical. This result holds when both variables are highly stable, in particular when development or change may be adequately represented by a single latent curve.

References

- Bass, B. M. (1962). Further evidence on the dynamic character of criteria. *Personnel Psychology, 15*, 93-97.
- Bentler, P. M. (1989). *EQS structural equations program manual*. Los Angeles: BMDP Statistical Software.
- Bock, R. D., & Bargmann, R. E. (1966). Analysis of covariance structures. *Psychometrika, 31*, 507-534.
- Cattell, R. B. (1964). Validity and reliability: A proposed more basic set of concepts. *Journal of Educational Psychology, 55*, 1-22.
- Collins, L. M. (1991). Measurement in longitudinal research. In L. M. Collins & J. L. Horn (Eds.), *Best methods for the analysis of change* (pp. 137-148). Washington DC: American Psychological Association.
- Collins, L. M., & Cliff, N. (1990). Using the longitudinal Guttman simplex as a basis for measuring growth. *Psychological Bulletin, 108*, 128-134.
- Collins, L. M., & Horn, J. L. (Eds.). (1991). *Best methods for the analysis of change*. Washington DC: American Psychological Association.
- Cronbach, L. J., Gleser, G. C., Nanda, H., & Rajaratnam, N. (1972). *The dependability of behavioral measurements: Theory of generalizability for scores and profiles*. New York: Wiley.
- Fleishman, E. A., & Fruchter, B. (1960). Factor structure and predictability of successive stages of learning Morse code. *Journal of Applied Psychology, 44*, 97-101.
- Ghiselli, E. E., & Haire, M. (1960). The validation of selection tests in the light of dynamic character of criteria. *Personnel Psychology, 13*, 225-231.
- Heise, D. R. (1969). Separating reliability and stability in test-retest correlation. *American Sociological Review, 34*, 93-101.
- Jöreskog, K. G. (1969). A general approach to confirmatory maximum likelihood factor analysis. *Psychometrika, 34*, 183-202.
- Jöreskog, K. G. (1971). Statistical analysis of sets of congeneric tests. *Psychometrika, 36*, 109-133.
- Jöreskog, K. G., & Sörbom, D. (1989). *LISREL 7: A guide to the program and applications* (2nd ed.). Chicago: SPSS, Inc.
- Lord, F. M., & Novick, M. R. (1968). *Statistical theories of mental test scores*. Reading MA: Addison-Wesley.
- Meier, S. T. (1993). Revitalizing the measurement curriculum. *American Psychologist, 48*, 886-891.
- Menard, S. (1991). *Longitudinal research*. (Sage University Paper Series on Quantitative Applications in the Social Sciences, 07-076). Newbury Park CA: Sage.
- Meredith, W., & Tisak, J. (1990). Latent curve analysis. *Psychometrika, 55*, 107-122.
- Millsap, R. E., & Everson, H. (1991). Confirmatory measurement model comparisons using latent means. *Multivariate Behavioral Research, 26*, 479-497.
- Morrison, D. F. (1990). *Multivariate statistical methods* (3rd ed.). New York: McGraw-Hill.
- Nesselroade, J. R., & Baltes, P. B. (1974). Adolescent personality development and historical change: 1970-1972. *Monographs of the Society for Research in Child Development, 39*, (1, Serial No. 154).
- Protter, M. H., & Morrey, C. B., Jr. (1970). *College calculus with analytic geometry* (2nd ed.). Reading MA: Addison-Wesley.
- Rao, C. R. (1958). Some statistical methods for comparison of growth curves. *Biometrics, 14*, 1-17.
- Rogosa, D. R., Brandt, D., & Zimowski, M. (1982). A growth curve approach to the measurement of change. *Psychological Bulletin, 90*, 726-748.

- Rogosa, D. R., & Willett, J. B. (1985). Understanding correlates of change by modeling individual differences in growth. *Psychometrika*, 50, 203–228.
- SAS Institute, Inc. (1989). *SAS/STAT user's guide* (Version 6, 4th ed., Vol. 1). Cary NC: Author.
- Schaie, K. W. (1965). A general model for the study of developmental problems. *Psychological Bulletin*, 64, 92–107.
- Scher, A. M., Young, A. C., & Meredith, W. M. (1960). Factor analysis of the electrocardiogram. *Circulation Research*, 8, 519–526.
- Summers, G. F., Hough, R. L., Scott, J. T., & Folse, C. L. (1969). *Before industrialization: A rural social system base study* (Bulletin No. 736). Urbana: Illinois Agricultural Experiment Station, University of Illinois.
- Thurstone, L. L., & Thurstone, T. G. (1962). *SRA Primary Mental Abilities*. Chicago: Science Research Associates.
- Tisak, J., & Meredith, W. (1990). Descriptive and associative developmental models. In A. von Eye (Ed.), *Statistical methods in developmental research. Volume II: Time series and categorical longitudinal data* (pp. 387–406). New York: Academic Press.
- Tucker, L. R. (1958). Determination of parameters of a functional relation by factor analysis. *Psychometrika*, 23, 19–23.
- Verdonik, F., & Sherrod, L. R. (1984). *An inventory of longitudinal research on childhood and adolescence*. New York: Social Science Research Council.
- von Eye, A., & Clogg, C. C. (1994). *Latent variables analysis*. Thousand Oaks CA: Sage.
- Wheaton, B., Muthén, B., Alwin, D. F., & Summers, G. F. (1977). Assessing reliability and stability in panel models. In D. R. Heise (Ed.), *Sociological methodology 1977* (pp. 84–136). San Francisco: Jossey Bass.
- Willett, J. B. (1989). Some results on reliability for the longitudinal measurement of change: Implications for the design of studies of individual growth. *Educational and Psychological Measurement*, 49, 587–602.

Acknowledgments

The authors gratefully acknowledge the assistance of Brett Plummer, Eugene F. Stone-Romero, and an anonymous reviewer for their critical comments and constructive suggestions on this manuscript.

Author's Address

Send requests for reprints or further information to John Tisak, Department of Psychology, Bowling Green State University, Bowling Green OH 43403-0228, U.S.A. Email: jtisak@trapper.bgsu.edu.