

# VU Research Portal

## The genetic analysis of repeated measures I. Simplex models

Boomsma, D.I.; Molenaar, P.C.M.

***published in***

Behavior Genetics  
1987

***DOI (link to publisher)***

[10.1007/BF01065991](https://doi.org/10.1007/BF01065991)

***document version***

Publisher's PDF, also known as Version of record

[Link to publication in VU Research Portal](#)

***citation for published version (APA)***

Boomsma, D. I., & Molenaar, P. C. M. (1987). The genetic analysis of repeated measures I. Simplex models. *Behavior Genetics*, 17(2), 111-123. <https://doi.org/10.1007/BF01065991>

**General rights**

Copyright and moral rights for the publications made accessible in the public portal are retained by the authors and/or other copyright owners and it is a condition of accessing publications that users recognise and abide by the legal requirements associated with these rights.

- Users may download and print one copy of any publication from the public portal for the purpose of private study or research.
- You may not further distribute the material or use it for any profit-making activity or commercial gain
- You may freely distribute the URL identifying the publication in the public portal ?

**Take down policy**

If you believe that this document breaches copyright please contact us providing details, and we will remove access to the work immediately and investigate your claim.

**E-mail address:**

[vuresearchportal.ub@vu.nl](mailto:vuresearchportal.ub@vu.nl)

## **The Genetic Analysis of Repeated Measures. I. Simplex Models**

**Dorret I. Boomsma<sup>1</sup> and Peter C. M. Molenaar<sup>2</sup>**

*Received 21 July 1985—Final 10 May 1986*

---

*The well-known simplex model is extended to a model that may be used for the genetic and environmental analysis of covariance structures. This "double" simplex structure can be specified as a LISREL model. It is shown that data which give rise to a simplex correlation structure, such as repeated-measures data, do not fit a factor-analysis model. The parameter estimation of the simplex model is illustrated with computer-simulated twin data.*

---

**KEY WORDS:** repeated measures; longitudinal data; simplex models; genetic correlations; environmental correlations; twin data; LISREL.

### **INTRODUCTION**

We define a repeated-measures study as a study in which the same variable or set of variables is measured repeatedly over time on the same subjects. The measurement occasions may be trials, such as in a learning experiment, or time periods, such as in a longitudinal study or in a psychophysiological experiment. The covariance or correlation matrix of such data will form a simplex (Guttman, 1954), that is, correlations will be maximal nearest the diagonal (i.e., among adjoining occasions) and fall away systematically as the distance between the variables along the time dimension increases. Stated otherwise, correlations correspond to distances between scale points.

A simplex structure can be generated by a Markov stochastic process (Jöreskog, 1970), that is, a first-order nonstationary autoregressive pro-

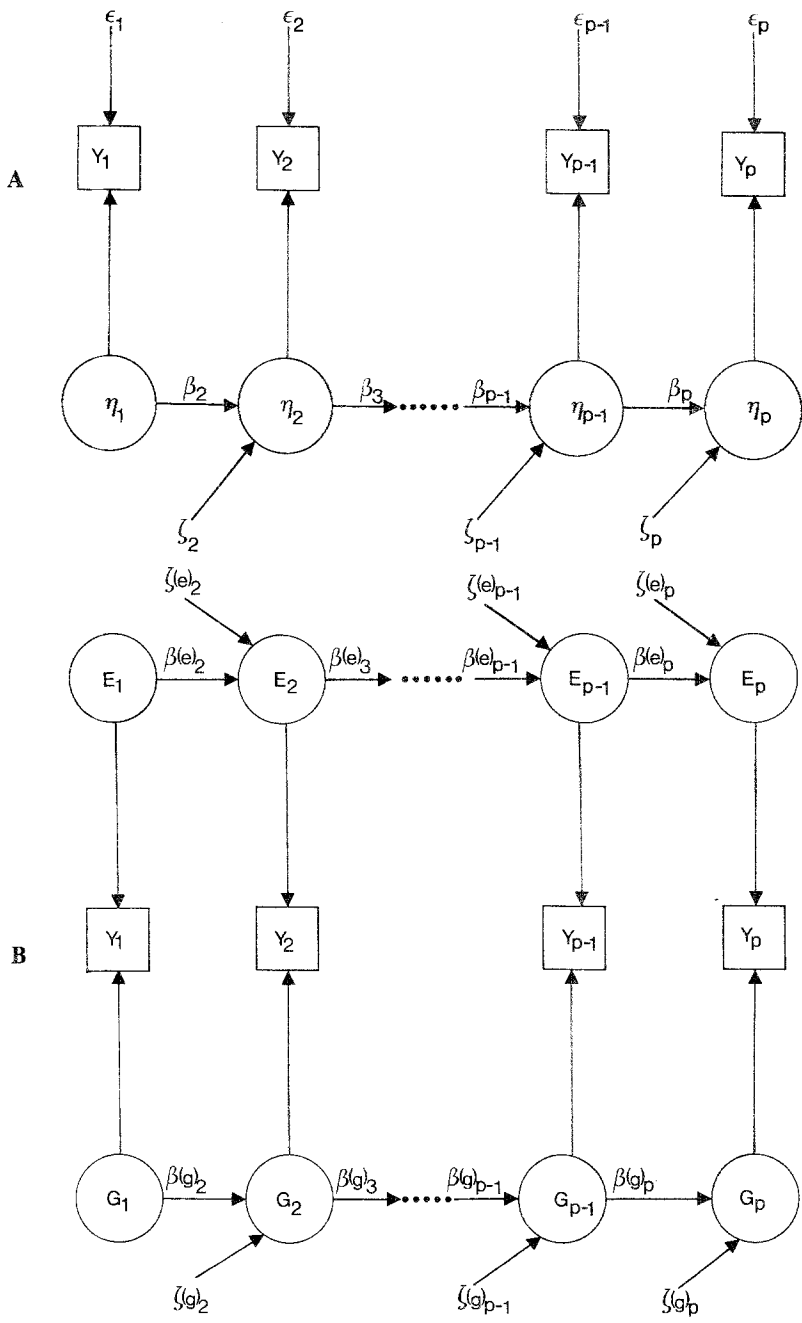
<sup>1</sup> Free University, Department of Experimental Psychology, De Boelelaan 1115, 1081 HV Amsterdam, The Netherlands.

<sup>2</sup> University of Amsterdam, Department of Psychology, Weesperplein 8, 1018 XA Amsterdam, The Netherlands.

cess. The Markov property of the model is that the partial correlation  $\rho_{ik \cdot j} = 0$ , whenever  $i < j < k$ . Guttman (1954) showed that the inverse of a perfect simplex correlation matrix, that is, a simplex with negligible measurement error, is of a very simple form: all elements are zero except for those in the three central diagonals. For a positive correlation matrix the main diagonal elements of its inverse are all positive; the elements in the two diagonals surrounding the main diagonal are negative and smaller than the elements on the main diagonal. From this it follows that, for prediction purposes, any intermediate variable can be predicted from two variables, i.e., the ones immediately before and after. To predict the first and last variables only the second and the next to the last variable, respectively, get nonzero regression weights. In the case of a quasi-simplex, which allows for substantial measurement errors, the multiple regression weights show the same general pattern. The highest weights usually go to the two neighbors of an intermediate variable. These two weights are always positive; all other weights are closer to zero and not restricted to be positive. According to Wold's decomposition theorem (cf. Hannan, 1977), this bidirectional prediction scheme is formally equivalent with a unidirectional prediction involving only the preceding variable (whence it is called a causal prediction).

Data that give rise to a simplex correlation structure will not usually fit a factor-analysis model with one common factor (Jöreskog, 1970). Furthermore, as pointed out by Cronbach (1967), factor analysis of any correlation matrix conforming to a simplex will almost always result in at least three or four factors whose loadings will increase and decrease with time according to a pattern expected from the mathematical properties of a simplex. Cronbach argued that there is nothing determinate in the points at which these factors peak or cross over, since these can be markedly altered by changing either the method of factor analysis or the time points included in the analysis. A more basic point was brought forward by Wohlwill (1973), namely, that the factor-analysis model is fundamentally unsuited to data conforming to a simplex, because of the determination of the correlations by the single dimension of proximity. In fact, Guttman (1954) already showed that for an infinite universe, a nonsingular simplex has an infinite number of common factors (p. 313).

For a repeated-measures design this implies that by simply increasing the number of time points in the analysis, a confirmatory factor analysis will result in an increasing number of factors before the model fits the data. Therefore, instead of a model based on factor analysis, we need an alternative model which recognizes the time-dependent structure of the data.



**Fig. 1.** (A) Quasi-simplex model, where  $y$ 's are the observed variables,  $\eta$ 's are the latent variables,  $\epsilon$ 's are measurement errors in  $y$ ,  $\beta$ 's are regression coefficients of  $\eta$  on  $\eta$ , and  $\zeta$ 's are residuals. (B) Double simplex model, where  $y$  is the observed time series,  $G$  and  $E$  are the underlying genetic and environmental series,  $\beta$ 's are regression coefficients, and  $\zeta$ 's are residuals.

### THE "SINGLE" SIMPLEX

Guttman formulated the idea of a simplex both with and without reference to a hypothetical underlying variable. We formulate here a model for a time series  $y_i$  with reference to an underlying time series  $\eta_i$ . If we consider  $p$  variables  $y_1, y_2, \dots, y_p$ , then a quasi-Markov simplex with  $p$  underlying latent variables  $\eta_1, \eta_2, \dots, \eta_p$ , can be represented in terms of path analysis as is shown in Fig. 1A. When the units of measurement in the latent variables are chosen to be same as in the observed variables, the equations defining the model are then

$$y_i = \eta_i + \epsilon_i, \quad i = 1, 2, \dots, p,$$

$$\eta_i = \beta_i \eta_{i-1} + \zeta_i, \quad i = 2, 3, \dots, p,$$

where the  $\epsilon$ 's are errors of measurement in  $y$  and the  $\zeta$ 's are random disturbance terms.

### THE "DOUBLE" SIMPLEX

Within a behavioral genetics context we usually work with more than one latent construct, i.e., we usually want to analyze phenotypic variance and covariance into at least genetic and environmental factors. We therefore need to extend the simplex model described above to a model in which more than a single underlying time series is specified. Such a model may be obtained as follows: consider  $p$  variables  $y_1, y_2, \dots, y_p$  (Fig. 1B), with latent variables  $G_1, G_2, \dots, G_p$  and  $E_1, E_2, \dots, E_p$ . The units of measurement in the latent variables are again chosen to be the same as in the observed variables. In this case, the equations defining the model become

$$y_i = G_i + E_i, \quad i = 1, 2, \dots, p,$$

$$G_i = \beta(g)_i G_{i-1} + \zeta(g)_i, \quad i = 2, 3, \dots, p,$$

$$E_i = \beta(e)_i E_{i-1} + \zeta(e)_i, \quad i = 2, 3, \dots, p,$$

where  $G_i$  and  $E_i$  and  $\zeta(g)_i$  and  $\zeta(e)_i$ ,  $i = 1, 2, \dots, p$ , are mutually uncorrelated and where  $\zeta(g)_{i+1}$  is uncorrelated with  $G_i$ , and  $\zeta(e)_{i+1}$  is uncorrelated with  $E_i$ ,  $i = 2, 3, \dots, p - 1$ .

The parameters of the model are

$$\Psi(g)_1 = \text{var}(G_1),$$

$$\Psi(e)_1 = \text{var}(E_1),$$

$$\Psi(g)_i = \text{var}[\zeta(g)_i], \quad i = 2, 3, \dots, p,$$

$$\Psi(e)_i = \text{var}[\zeta(e)_i], \quad i = 2, 3, \dots, p,$$

$$\beta(g)_2, \beta(g)_3, \dots, \beta(g)_p, \quad \text{and} \quad \beta(e)_2, \beta(e)_3, \dots, \beta(e)_p.$$

If  $p = 3$  the covariance matrix of  $y_i$  is of the form

$$\begin{bmatrix} \text{var}(G_1) + \text{var}(E_1) & & & \\ [\beta(g)_2 \text{var}(G_1) & \text{var}(G_2) + \text{var}(E_2) & & \\ + \beta(e)_2 \text{var}(E_1)] & & & \\ [\beta(g)_2 \beta(g)_3 \text{var}(G_1) & [\beta(g)_3 \text{var}(G_2) & \text{var}(G_3) + \text{var}(E_3) \\ + \beta(e)_2 \beta(e)_3 \text{var}(E_1)] & + \beta(e)_3 \text{var}(E_2)] & & \end{bmatrix}$$

where

$$\text{var}(G_i) = \beta^2(g)_i \text{var}(G_{i-1}) + \Psi(g)_i, \quad i = 2, 3, \dots, p,$$

$$\text{var}(E_i) = \beta^2(e)_i \text{var}(E_{i-1}) + \Psi(e)_i, \quad i = 2, 3, \dots, p.$$

With this model no measurement errors are specified. Any measurement errors in  $y$  are included in the separate environmental factors. The identification of this model, however, requires data from genetically related individuals. We illustrate the model and the parameter estimation with computer-simulated twin data, but these data can be replaced by or extended to other family groupings as well.

### DATA SIMULATION

Time series were generated for 100 monozygotic (MZ) and 100 dizygotic (DZ) twin pairs using the simple genetic model  $y_i = G_i + E_i$ ,  $i = 1, 2, \dots, 10$ , under the assumptions that gene action is additive, mating is random, and there is no shared environmental influence. For both  $G_i$  and  $E_i$  a first-order autoregressive model was specified:

$$G_i = \beta(g)_i G_{i-1} + \zeta(g)_i, \quad i = 2, 3, \dots, 10,$$

$$E_i = \beta(e)_i E_{i-1} + \zeta(e)_i, \quad i = 2, 3, \dots, 10.$$

With respect to  $G$  and  $E$  at each time point,

$$\text{cor} [E_{1i}, E_{2i}] = 0,$$

$$\text{cor} [G_{1i}, G_{2i}] = 1 \text{ for MZ twins, and}$$

$$\text{cor} [G_{1i}, G_{2i}] = 0.5 \text{ for DZ twins,}$$

where the first subscript identifies persons within twin pairs.

Using the FTGEN subroutine from the IMSL Fortran Library (IMSL, Inc., 1979), four MZ and DZ data sets were generated:

(I)  $\beta(g)_i = 0$       and       $\beta(e)_i = 0$ ;

(II)  $\beta(g)_i = 0.75$       and       $\beta(e)_i = 0$ ;

Table I. Simplex Correlation Matrix (Data IV)

1									
0.693	1								
0.438	0.736	1							
0.345	0.605	0.763	1						
0.247	0.443	0.567	0.733	1					
0.179	0.302	0.413	0.527	0.745	1				
0.153	0.217	0.257	0.351	0.532	0.765	1			
0.171	0.232	0.249	0.293	0.406	0.561	0.769	1		
0.130	0.163	0.210	0.250	0.345	0.449	0.541	0.737	1	
0.127	0.175	0.174	0.186	0.247	0.291	0.413	0.579	0.743	1

(III)  $\beta(g)_i = 0$  and  $\beta(e)_i = 0.75$ ;

(IV)  $\beta(g)_i = 0.75$  and  $\beta(e)_i = 0.75$ .

All  $G_i$  and  $E_i$  were standardized with mean 0 and variance 100 across all time points, whence the proportion of genetic variance at each time point is always fixed at 0.5. If  $\beta(g)_i \neq 0$  or  $\beta(e)_i \neq 0$ , then the covariance function of  $G_i$  or  $E_i$  is a decreasing function of the distance between time points, whence the correlation matrix of the observed time series  $y_i$  conforms to a simplex structure. For the last data set (data IV) the simplex structure of the resulting correlation matrix is shown in Table I.

### DATA ANALYSIS

As a first step between-pair and within-pair mean products matrices were computed for MZ and DZ twins for each of the four data sets. Each  $10 \times 10$  matrix has 55 unique statistics, providing a total of 220 df. All data sets were analyzed with both confirmatory factor-analysis models (Martin and Eaves, 1977; Boomsma and Molenaar, 1986) and simplex models. For the factor-analysis model the expected covariance structures for the MZ and DZ between and within matrices may be written as follows:

$$\Sigma_{\text{MZB}} = 2(\Delta\Delta' + \mathbf{D}^2) + \mathbf{H}\mathbf{H}' + \mathbf{E}^2,$$

$$\Sigma_{\text{MZW}} = \mathbf{H}\mathbf{H}' + \mathbf{E}^2,$$

$$\Sigma_{\text{DZB}} = 1.5(\Delta\Delta' + \mathbf{D}^2) + \mathbf{H}\mathbf{H}' + \mathbf{E}^2,$$

$$\Sigma_{\text{DZW}} = 0.5(\Delta\Delta' + \mathbf{D}^2) + \mathbf{H}\mathbf{H}' + \mathbf{E}^2,$$

where  $\Delta$  represents the loadings of the variables on the additive genetic factors, and  $\mathbf{H}$  the loadings on the environmental factors.  $\mathbf{D}$  and  $\mathbf{E}$  represent diagonal matrices containing loadings of genetic and environmental

influences specific to each variable. The computational procedures for this model using the LISREL computer program have been described elsewhere (Boomsma and Molenaar, 1986).

For the simplex models the expected covariance structures can be written as

$$\begin{aligned}\Sigma_{MZB} &= 2(\mathbf{B}_g \mathbf{G} \mathbf{B}_g') + \mathbf{B}_e \mathbf{E} \mathbf{B}_e', \\ \Sigma_{MZW} &= \mathbf{B}_e \mathbf{E} \mathbf{B}_e', \\ \Sigma_{DZB} &= 1.5(\mathbf{B}_g \mathbf{G} \mathbf{B}_g') + \mathbf{B}_e \mathbf{E} \mathbf{B}_e', \\ \Sigma_{DZW} &= 0.5(\mathbf{B}_g \mathbf{G} \mathbf{B}_g') + \mathbf{B}_e \mathbf{E} \mathbf{B}_e',\end{aligned}$$

where  $\mathbf{G}$  and  $\mathbf{E}$  are  $(p \times p)$  diagonal matrices containing the unique genetic and environmental variances at each time point and where  $\mathbf{B}_g$  and  $\mathbf{B}_e$  ( $p \times p$ ) are coefficient matrices with the same form. For instance,

$$\mathbf{B}_g = \begin{bmatrix} 1 & & & & & & \\ \beta(g)_2 & & & & & & \\ \beta(g)_2\beta(g)_3 & & 1 & & & & \\ \cdot & & \beta(g)_3 & & 1 & & \\ \cdot & & \cdot & & \cdot & & 1 \\ \beta(g)_2 \cdots \beta(g)_p & & & & & & \beta(g)_p & & 1 \end{bmatrix}.$$

Notice that  $\mathbf{B}_g$  can be written as

$$\mathbf{B}_g = (\mathbf{I} - \mathbf{B}_g^*)^{-1},$$

where  $\mathbf{B}_g^*$  is defined as

$$\mathbf{B}_g^* = \begin{bmatrix} 0 & 0 & 0 & 0 & 0 & 0 & \\ \beta(g)_2 & 0 & 0 & 0 & 0 & 0 & \\ 0 & \beta(g)_3 & 0 & 0 & 0 & 0 & \\ 0 & 0 & 0 & \cdot & 0 & 0 & \\ \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \\ 0 & 0 & 0 & 0 & \beta(g)_p & 0 & \end{bmatrix}.$$

Likewise,  $\mathbf{B}_e$  may be written as  $\mathbf{B}_e = (\mathbf{I} - \mathbf{B}_e^*)^{-1}$ . The latter expressions, i.e.,  $\mathbf{B}_g^*$  and  $\mathbf{B}_e^*$ , will be used in order to arrive at a LISREL model (Jöreskog and Sörbom, 1981) for the proposed analysis.

In LISREL the covariance matrix of  $y$  equals

$$\Sigma_y = \Lambda_y (\mathbf{I} - \mathbf{B})^{-1} \Psi (\mathbf{I} - \mathbf{B}')^{-1} \Lambda_y'.$$

If the units of measurement in the latent variables are chosen to be the



same as in the observed variables, the  $\Lambda_y$  matrix can be used for the genetic and environmental weights from the model specified above. For  $p = 3$  the accommodation of the expected covariance structures for MZ twins can be obtained by

$$\begin{array}{l} \text{MZB,} \\ \Lambda_y = \begin{bmatrix} \sqrt{2} & 0 & 0 & 1 & 0 & 0 \\ 0 & \sqrt{2} & 0 & 0 & 1 & 0 \\ 0 & 0 & \sqrt{2} & 0 & 0 & 1 \end{bmatrix}; \end{array} \quad \begin{array}{l} \text{MZW,} \\ \Lambda_y = \begin{bmatrix} 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 \end{bmatrix}. \end{array}$$

For DZ twins the  $\Lambda_y$  matrices become

$$\begin{array}{l} \text{DZB,} \\ \Lambda_y = \begin{bmatrix} \sqrt{1.5} & 0 & 0 & 1 & 0 & 0 \\ 0 & \sqrt{1.5} & 0 & 0 & 1 & 0 \\ 0 & 0 & \sqrt{1.5} & 0 & 0 & 1 \end{bmatrix}; \end{array} \quad \begin{array}{l} \text{DZW,} \\ \Lambda_y = \begin{bmatrix} \sqrt{.5} & 0 & 0 & 1 & 0 & 0 \\ 0 & \sqrt{.5} & 0 & 0 & 1 & 0 \\ 0 & 0 & \sqrt{.5} & 0 & 0 & 1 \end{bmatrix}. \end{array}$$

Note that the weighting of  $(\mathbf{B}_g \mathbf{G} \mathbf{B}_g')$  in the expected covariance structures is accommodated by the square roots of the genetic weights in  $\Lambda_y$ , since the  $\Lambda_y$  matrix pre- and postmultiplies  $[(\mathbf{I} - \mathbf{B})^{-1} \Psi (\mathbf{I} - \mathbf{B}')^{-1}]$  in the construction of  $\Sigma_y$ .

$\mathbf{B}$  ( $6 \times 6$ ) is composed of  $\mathbf{B}_g^*$  and  $\mathbf{B}_e^*$ :

$$\mathbf{B} = \begin{bmatrix} 0 & 0 & 0 & 0 & 0 & 0 \\ \beta(g)_2 & 0 & 0 & 0 & 0 & 0 \\ 0 & \beta(g)_3 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & \beta(e)_2 & 0 & 0 \\ 0 & 0 & 0 & 0 & \beta(e)_3 & 0 \end{bmatrix},$$

and  $\Psi$  ( $6 \times 6$ ) is a diagonal matrix containing the unique genetic and environmental variances, i.e., that part of the variances that is not explained by  $G_{i-1}$  or  $E_{i-1}$ .

From these specifications it follows that  $\Sigma_y =$

$$\begin{bmatrix} \sigma_{11} & & & & & \\ \sigma_{12} & \sigma_{22} & & & & \\ \sigma_{13} & \sigma_{23} & \sigma_{33} & & & \end{bmatrix},$$

where

$$\sigma_{11} = \Psi(g)_1 + \Psi(e)_1,$$

$$\sigma_{12} = \beta(g)_2\Psi(g)_1 + \beta(e)_2\Psi(e)_1,$$

$$\sigma_{22} = \beta^2(g)_2\Psi(g)_1 + \Psi(g)_2 + \beta^2(e)_2\Psi(e)_1 + \Psi(e)_2,$$

$$\sigma_{13} = \beta(g)_2\beta(g)_3\Psi(g)_1 + \beta(e)_2\beta(e)_3\Psi(e)_1,$$

$$\sigma_{23} = \beta(g)_3[\beta^2(g)_2\Psi(g)_1 + \Psi(g)_2] + \beta(e)_3[\beta^2(e)_2\Psi(e)_1 + \Psi(e)_2],$$

$$\begin{aligned} \sigma_{33} = & \beta^2(g)_3[\beta^2(g)_2\Psi(g)_1 + \Psi(g)_2] + \Psi(g)_3 \\ & + \beta^2(e)_3[\beta^2(e)_2\Psi(e)_1 + \Psi(e)_2] + \Psi(e)_3. \end{aligned}$$

The four equations  $\sum_y = \Lambda_y (\mathbf{I} - \mathbf{B})^{-1} \Psi (\mathbf{I} - \mathbf{B}')^{-1} \Lambda_y'$  can be fit simultaneously with LISREL using a four-group design with parameters to be estimated invariant across groups. In addition to estimates of all  $\beta$ 's and  $\Psi$ 's, the standardized LISREL solution also gives the matrices of genetic and environmental correlations between time points.

## RESULTS

Table II shows the results of fitting a factor-analysis model with one common genetic and one common environmental factor and a simplex model as outlined in the preceding section to all four data sets.

For the first data set [where  $\beta(g)_i = \beta(e)_i = 0$ ] the  $\chi^2$ 's for both the factor-analysis model and the simplex model indicate a good fit. Both models correctly recover that there is no genetic or environmental covariance among time points. In the factor-analysis model this was indicated by the nonsignificant loadings on the common factors and in the simplex model by the nonsignificant regression weights.

Table II. Model Fitting

	Factor-analysis model (df = 180)		Simplex model (df = 182)	
	$\chi^2$	<i>P</i>	$\chi^2$	<i>P</i>
Data I	170.6	0.68	182.4	0.48
Data II	289.2	0.00	179.1	0.55
Data III	658.3	0.00	195.3	0.24
Data IV	1140.8	0.00	204.2	0.13

Table III. Model Fitting (Data IV)

	df	$\chi^2$	<i>P</i>
Simplex model			
	182	204.23	0.13
Factor-analysis models			
1 <i>G</i> + 1 <i>E</i>	180	1140.79	0.00
2 <i>G</i> + 2 <i>E</i>	162	478.92	0.00
3 <i>G</i> + 3 <i>E</i>	146	298.78	0.00
4 <i>G</i> + 4 <i>E</i>	132	169.31	0.02
5 <i>G</i> + 5 <i>E</i>	120	130.58	0.24

For the second [where all  $\beta(g) = 0.75$ ] and the third [where all  $\beta(e) = 0.75$ ] data sets, the factor-analysis models no longer fit the data, while the simplex models correctly identify the underlying structures of the genetic and environmental series. As indicated by the degrees of freedom in Table II, no parameters were fixed at zero in advance.

We now consider the last data set [where  $\beta(g)_i = \beta(e)_i = 0.75$ ] in some more detail. The factor-analysis model with one common genetic and one common environmental factor gave a very large  $\chi^2$  for this data set. Next, we systematically increased the number of common factors as is shown in Table III. For these models Cholesky factors were used, that is, a lower triangular matrix of factor loadings. This pattern conforms to a specific rotation in an otherwise unrestricted vector space (cf. Mulaik, 1972). For a time series consisting of 10 time points a total of five common genetic and five common environmental factors has to be used before the model fits the data. Moreover, the significant loadings on these common factors did not allow a meaningful interpretation of the data. In contrast, the simplex model gives not only a satisfactory fit, but also an adequate account of the underlying genetic and environmental covariance functions. LISREL parameter estimates for the simplex model are presented in Table IV. Also in Table IV are the estimates for the total amounts of genetic and environmental variance at each time point. From these variances we can summarize the contributions of the unique and shared genetic and environmental variances at each time point to the total variances as is done in the lower part of Table IV. As can be seen, roughly half of the genetic and environmental variances at each time point are unique in the sense that they are not explained by the preceding *G* or *E*, respectively.

Finally, Table V shows the matrices of genetic and environmental correlations among time points as they were obtained from the standardized LISREL solution. It is clear that each matrix conforms to a perfect simplex structure.

Table IV. Simplex Analysis (Data IV)<sup>a</sup>

	$\Psi(g)$	$\beta(g)$	$\text{var}(G)$	$\Psi(e)$	$\beta(e)$	$\text{var}(E)$	$\text{var}(T)$
1	98.0 (16.3)	—	98.0	87.1 (11.3)	—	87.1	185.1
2	44.5 (8.3)	0.65 (0.08)	85.9	49.7 (6.3)	0.69 (0.07)	91.2	177.1
3	46.1 (8.0)	0.82 (0.09)	103.9	43.5 (5.6)	0.73 (0.06)	92.1	196.0
4	46.7 (7.4)	0.78 (0.08)	109.9	37.4 (4.9)	0.77 (0.06)	92.0	201.9
5	44.8 (8.1)	0.82 (0.08)	118.7	44.5 (5.7)	0.64 (0.06)	82.2	200.9
6	52.8 (8.4)	0.81 (0.07)	130.7	42.9 (5.6)	0.77 (0.07)	91.6	222.3
7	41.9 (7.7)	0.66 (0.07)	98.8	41.9 (5.4)	0.87 (0.06)	111.2	210.0
8	41.6 (6.9)	0.77 (0.08)	117.7	36.8 (4.8)	0.71 (0.05)	92.9	210.6
9	51.6 (7.4)	0.71 (0.08)	110.9	33.0 (4.9)	0.76 (0.06)	86.7	197.6
10	45.1 (7.6)	0.65 (0.08)	92.0	37.6 (4.9)	0.85 (0.06)	100.2	192.2

## Proportions of variance

	Unique genetic	Total genetic	Unique environmental	Total environmental
1	—	0.53	—	0.47
2	0.25	0.49	0.28	0.51
3	0.24	0.53	0.22	0.47
4	0.23	0.54	0.19	0.46
5	0.22	0.59	0.22	0.41
6	0.24	0.59	0.19	0.41
7	0.20	0.47	0.20	0.53
8	0.20	0.56	0.18	0.44
9	0.26	0.56	0.17	0.44
10	0.24	0.48	0.20	0.52

<sup>a</sup>  $\Psi(g)$ ,  $\beta(g)$ ,  $\Psi(e)$ , and  $\beta(e)$  are LISREL parameter estimates (standard errors are given in parentheses).  $\text{var}(G_i) = \beta^2(g)_i \text{var}(G_{i-1}) + \Psi(g)_i$ ,  $\text{var}(E_i) = \beta^2(e)_i \text{var}(e_{i-1}) + \Psi(e)_i$ , and  $\text{var}(T) = \text{var}(G) + \text{var}(E)$ .

## DISCUSSION

The above results clearly indicate the failure of factor-analysis models in analyzing data that conform to a simplex structure, such as repeated-measures data. More generally, multivariate methods that have been developed for the analysis of covariance of different types of variables should not be generalized directly to longitudinal data. In fact, a simplex

Table V. Genetic and Environmental Correlations

Genetic Correlations (data IV)										
1	1									
2	0.70	1								
3	0.52	0.75	1							
4	0.39	0.57	0.76	1						
5	0.31	0.45	0.60	0.79	1					
6	0.24	0.34	0.46	0.61	0.77	1				
7	0.18	0.26	0.35	0.46	0.58	0.76	1			
8	0.14	0.20	0.27	0.35	0.45	0.59	0.76	1		
9	0.10	0.14	0.19	0.25	0.31	0.40	0.54	0.70	1	
10	0.07	0.10	0.13	0.17	0.22	0.29	0.37	0.49	0.70	1
Environmental correlations										
1	1									
2	0.67	1								
3	0.49	0.73	1							
4	0.38	0.56	0.77	1						
5	0.25	0.38	0.52	0.68	1					
6	0.18	0.27	0.38	0.49	0.73	1				
7	0.15	0.22	0.30	0.39	0.57	0.79	1			
8	0.11	0.17	0.23	0.30	0.45	0.61	0.78	1		
9	0.09	0.13	0.18	0.24	0.35	0.48	0.61	0.79	1	
10	0.07	0.10	0.14	0.19	0.28	0.38	0.48	0.62	0.79	1

structure may even apply to variables that are ordered not in time, but along a dimension of complexity as discussed by, for instance, Guttman (1954) and Jöreskog (1970).

We showed that the simplex model can be extended to the analysis of more than a single latent construct, so that it becomes possible to analyze an observed time series into underlying genetic and environmental series. In contrast with Jöreskog's (1970) original model, no specific constraints are required to make the model identified, since this generalized model includes separate simplex structures for  $G_i$  and  $E_i$ , and measurement errors are included in  $E$ .

Although in the illustrative models above the proportions of genetic and environmental variance were equal across time and all genetic and environmental correlations between time points were the same, this is not a necessary restriction. In fact, no such constraints were used in the LISREL analysis. Thus, the simplex model allows for both differential heritabilities and environmentalities at different time points and different genetic and environmental correlations between time points. Also, the assumption of no shared family environment is not a necessary restriction.

An additional structure may be specified to test the influence of common-environmental components.

The proposed simplex analysis can also be used for the analysis of multivariate time series. In this case, each single time series may conform to a simplex structure, while the relationship between different types of variables can be analyzed with a confirmatory factor-analysis model. All these analyses can be carried out with LISREL. There is a limit, however, to the number of variables that can be used in a LISREL analysis. We therefore developed an alternative approach to the genetic analysis of time series which will be presented in a subsequent paper.

### REFERENCES

- Boomsma, D. I., and Molenaar, P. C. M. (1986). Using LISREL to analyze genetic and environmental covariance structure. *Behav. Genet.* **16**:237–250.
- Cronbach, L. J. (1967). Year-to-year correlations of mental tests: A review of the Hofstaetter analysis. *Child Dev.* **38**:283–289.
- Guttman, L. (1954). A new approach to factor analysis: the radex. In Lazarsfeld, P. F. (ed.), *Mathematical Thinking in the Social Sciences*, Free Press, Glencoe, Ill., pp. 258–349.
- Hannan, E. J. (1977). *Multiple Time Series*, Wiley, New York.
- IMSL, Inc. (1979). *IMSL Library Reference Manual Edition 7*, IMSL Inc., Houston, Tex.
- Jöreskog, K. G. (1970). Estimation and testing of simplex models. *Br. J. Math. Stat. Psychol.* **23**:121–145.
- Jöreskog, K. G., and Sörbom, D. (1981). *LISREL: Analysis of Linear Structural Relationships by Maximum Likelihood and Least Squares Methods*, National Educational Resources, Chicago.
- Martin, N. G., and Eaves, L. J. (1977). The genetical analysis of covariance structure. *Heredity* **38**:79–95.
- Mulaik, S. A. (1972). *The Foundations of Factor Analysis*, McGraw–Hill, New York.
- Wohlwill, J. F. (1973). *The Study of Behavioral Development*, Academic Press, New York.

Edited by C. R. Cloninger