DOCUMENT RESUME

ED 259 007 TM 850 376

AUTHOR	Hanna, Gila; Lei, Hau
TITLE	A Longitudinal Analysis Using the LISREL-Model with Structured Means.
PUB DATE	Apr 84
NOTE	13p.; Paper presented at the Annual Meeting of the American Educational Research Association (68th, New Orleans, LA, April 23-27, 1984).
PUB TYPE	Speeches/Conference Papers (150) -- Reports - Research/Technical (143)
EDRS PRICE	MF01/PC01 Plus Postage.
DESCRIPTORS	Analysis of Covariance; Computer Software; Elementary School Mathematics; *Evaluation Methods; Factor Analysis; Foreign Countries; French; Goodness of Fit; Immersion Programs; Intermediate Grades; *Longitudinal Studies; *Mathematical Models; *Mathematics Achievement; *Program Evaluation; Second Language Programs; Statistical Analysis
IDENTIFIERS	Canadian Tests of Basic Skills; *LISREL Computer Program; Ontario; *Structural Equation Models

ABSTRACT

This paper discusses the use of the LISREL computer program to compare two groups of Ontario students who studied French as a second language in either a regular or an immersion program. Longitudinal data on growth in mathematical ability, from grade 4 to grade 6, are analyzed using seven different mathematical models. The paper illustrates the use of LISREL for models with structured means, and demonstrates its usefulness for investigating the structural relationships among measurements of one true variable taken at yearly intervals. (GDC)

A LONGITUDINAL ANALYSIS USING THE LISREL-MODEL WITH STRUCTURED MEANS

Gila Hanna and Hau Lei

Ontario Institute for Studies in Education

Paper presented at the Annual Meeting of the American Educational Research Association

New Orleans, April 1984

A LONGITUDINAL ANALYSIS USING
THE LISREL-MODEL WITH STRUCTURED MEANS

G. Hanna and H. Lei
Ontario Institute for Studies in Education

This paper reports on the use of LISREL to compare two groups of students, each group following a different program, from the point of view of the growth in mathematical ability from Grade 4 to Grade 6. The main purpose of the paper is to illustrate the use of the LISREL program for models with structured means, and in particular to demonstrate its usefulness for investigating the structural relationships among measurements of one true variable taken at yearly intervals.

To investigate the structural relationships among the successive measurements of the true variable postul...ed in the present study, a series of three-wave two-variable models was developed using techniques drawn from a number of existing factor-analytic models: the multi-wave multi-variable model (Jöreskog, 1979), the simultaneous factor-analysis model in several populations (Jöreskog, 1971), and the structural equation models with structured means (Sörbom, 1974, 1978). The latter model is of particular relevance to this study though the examples discussed by Sörbom (1978, 1982) all deal with a single structural equation and a single latent covariate. This study expands the application of the structural equation model with structured means to the case of two structural equations, one of which contains two latent covariates.

Each of the models developed in the present study was tested for goodness of fit by examining the chi-square statistic, the residuals, and the difference in chi-square values between the successive models of the series (Bentler, 1982; Jöreskog, 1979).

Description of the Data

The data used in this study were drawn from the data pool of the Bilingual Education Project. Data on achievement in English as a first language, French as a second language, and academic subjects were collected in the course of a large-scale evaluation of Boards of Education in Ontario which had both Immersion and Regular programs. Students enrolled in a French Immersion program get their instruction in French, their second language, while students enrolled in the Regular program get their instruction in English, their first language. According to Genessee (1983) the goals of the Immersion program are:

> (1) to provide the participating students with functional competence in the second language which may or may not be native-like; (2) to maintain and develop normal levels of first-language competence; (3) to ensure achievement in academic subjects commensurate with the students' academic ability at grade level; and (4) to instill in the students an understanding and appreciation for the target-language group, their language and culture without detracting in any way from the students' identity with and appreciation for the home language and culture.

From 1970 to 1979 yearly evaluations were carried out to assess the degree to which the goals of the Immersion program were attained and to compare the academic achievement of students in the Immersion program to that of students in the Regular program. The results of these evaluations are summarized in Swain and Lapkin (1981).

The present study departs from the initial evaluations in two ways. First, the initial evaluations were cross-sectional, while this study is concerned with a longitudinal set of data. Second, the initial evaluations consisted of comparisons on the observed variables through univariate analysis of variance, while this study makes use of an analysis of covariance structures to analyse the underlying true variables.

For the present study a longitudinal set of data was extracted from the existing pool, encompassing the 144 Immersion students and 59 Regular students for whom there were complete results for the Canadian Test of Basic Skills (CTBS) tests in Mathematics Concepts and Mathematics Problems at three successive grade levels (Grades 4, 5, and 6). (There are fewer Regular than Immersion students in the longitudinal data because of the focus of the study for which the data were originally collected. When the data were collected the same Immersion students were tracked from year to year, while the composition of the Regular groups changed from year to year.)

Each student had six test scores, two at each grade level; all tests were administered in English. The two tests were as follows:

1. Mathematics Concepts: a test of about 30 multiple-choice items designed to test how well the student understands the number system and the terms and operations used in mathematics.

2. Mathematics Problems: a test of about 30 multiple-choice items designed to assess the student's skills in solving mathematical problems, using single-step or multiple-step problems.

The individual student scores are grade equivalent. In other words, the population mean score for Grade 4 was normalized to 40 at the beginning of the school year, to 45 five months later (the end of January), and to 50 at the end of June (or the beginning of Grade 5). Similarly the mean is 55 in the middle of Grade 5 and 65 in the middle of Grade 6. Since the tests were administered in April, the eighth month of the school year, the mean scores would be expected to be 48, 58, and 68 in Grades 4, 5, and 6, respectively.

Method

The Measurement Model

As shown in Figure 1, at each grade level the two tests, Mathematics Concepts and Mathematics Problems, are taken to be indicators of the latent variable η (mathematical ability); η_4, η_5, and η_6 represent the latent variable "mathematical ability" in Grades 4, 5, and 6, respectively.

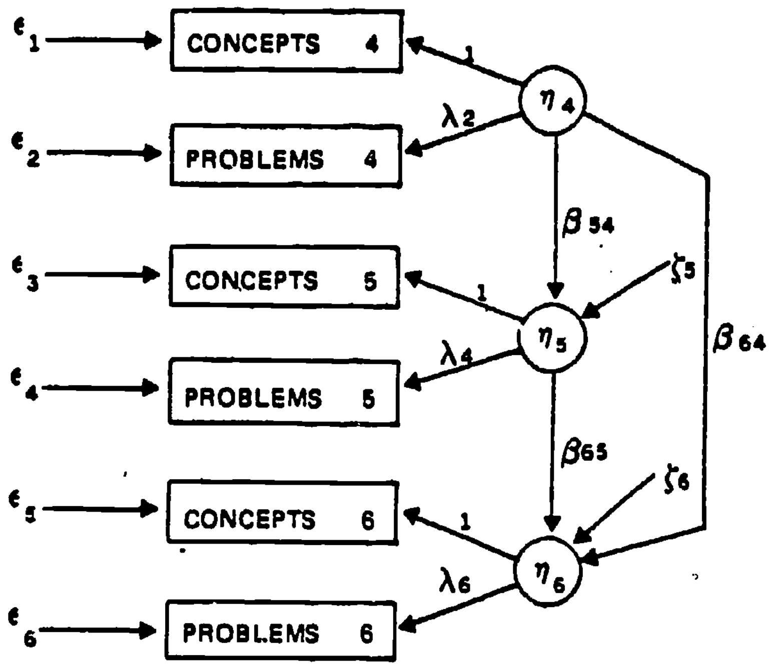

Figure 1: Model for the measurement
of growth in mathematical ability

The measurement model for mathematical ability is simply

$$\begin{bmatrix} y_5 \\ y_6 \\ y_3 \\ y_4 \\ y_1 \\ y_2 \end{bmatrix} = \begin{bmatrix} 1 & & & \nu_5 \\ \lambda_6 & & & \nu_6 \\ & 1 & & \nu_3 \\ & \lambda_4 & & \nu_4 \\ & & 1 & \nu_1 \\ & & \lambda_2 & \nu_2 \end{bmatrix} \begin{bmatrix} \eta_6 \\ \eta_5 \\ \eta_4 \\ 1 \end{bmatrix} + \begin{bmatrix} \varepsilon_5 \\ \varepsilon_6 \\ \varepsilon_3 \\ \varepsilon_4 \\ \varepsilon_1 \\ \varepsilon_2 \end{bmatrix}$$

The measurement equation assumes that the two tests, Mathematics Concepts and Mathematics Problems, measure a single latent variable on each occasion. The scale for each latent variable was set to be that of the test Mathematics Concepts, by fixing the appropriate factor loading equal to one for both groups. Also the measurement properties of the tests were constrained to be equal across the two groups, that is, they had the same origin of measurement (ν_i), the same loadings (λ_i), and the same error variances (ε_i), ($i = 1, 2, \ldots, 6$).

The Structural Model

The structural model is represented by the equation:

$$\begin{bmatrix} \eta_6 \\ \eta_5 \\ \eta_4 \\ 1 \end{bmatrix} = \begin{bmatrix} 0 & \beta_{65} & \beta_{64} & 0 \\ 0 & 0 & \beta_{54} & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix} \begin{bmatrix} \eta_6 \\ \eta_5 \\ \eta_4 \\ 1 \end{bmatrix} + \begin{bmatrix} \alpha_6 \\ \alpha_5 \\ \kappa_4 \\ 1 \end{bmatrix} [1] + \begin{bmatrix} \zeta_6 \\ \zeta_5 \\ \eta_4-\kappa_4 \\ 0 \end{bmatrix}$$

The structural model specifies a causal relationship among the true variables representing mathematical ability at Grades 4, 5, and 6. Specifically, it implies that η_4 has a direct effect on both η_5 and η_6, and that η_5 has an effect only on η_6.

It should be added that a more restricted quasi-simplex model (Joreskog, 1979), one in which there is no direct effect of η_4 on η_6 (i.e., $\beta_{64} = 0$), but there is an indirect effect mediated through η_5, was also postulated.

The origin of measurement and the mean of the latent variable cannot be identified simultaneously, but the differences between the groups can be estimated. When the mean of the latent variable, η_4, is fixed to zero for the Immersion group, the parameter representing the expectation of η_4 for the Regular group is the mean difference in initial mathematical ability between the two groups. The same reasoning applies to η_5 and η_6; only differences between the groups can be estimated. However, it is meaningful to consider these differences as the effect of program when β parameters (the regression weights) can be shown to be equal in the two groups.

Procedure

Following Jöreskog and Sörbom (1981), the present analysis of mean structures was carried out on an augmented (i.e., a constant variable 1 is added) moment matrix rather than on a covariance matrix. The input to the LISREL program (Tables 1 and 2) consisted of a correlation matrix, a vector of standard deviations, and a vector of means.

Table 1

Correlations, Means and Standard Deviat
Immersion Group N=144

Test	C4	P4	C5	P5	C6	P6	Means	Standard Deviations
Concepts 4	1						49.20	9.15
Problems 4	.753	1					48.10	9.53
Concepts 5	.733	.685	1				61.33	9.56
Problems 5	.634	.636	.705	1			58.61	10.32
Concepts 6	.715	.624	.726	.688	1		73.24	11.54
Problems 6	.603	.573	.669	.639	.729	1	67.37	11.60

Table 2

Correlations, Means and Standard Deviations
Regular Group N=59

Test	C4	P4	C5	P5	C6	P6	Means	Standard Deviations
Concepts 4	1						49.24	9.40
Problems 4	.657	1					46.58	9.30
Concepts 5	.631	.524	1				58.47	9.82
Problems 5	.637	.656	.614	1			55.14	10.86
Concepts 6	.617	.553	.716	.638	1		70.76	12.14
Problems 6	.494	.586	.486	.612	.672	1	68.08	10.81

The differences between the two groups were investigated by using a series of models. An initial model was first postulated, in which the measurement parameters were constrained to be invariant across the two groups but the structural parameters were left free in both groups. After verifying that this initial model yielded an acceptable fit, a series of nested models was tested, each new model being more restricted than the previous one in that one additional parameter is constrained.

Testing nested models makes it possible to assess the contribution of an individual parameter to the goodness of fit, through the following procedure. If A is a model with X degrees of freedom, a new model B may be formulated such that one additional parameter is either fixed to the value zero or constrained to be invariant across the two groups. B, then, has X-1 degrees of freedom. The new restriction can then be tested by looking at the chi-square difference (A-B) with 1 degree of freedom. If the difference between the chi square for A and that for B does not exceed the critical value at a chosen level, then the new restriction cannot be rejected. If, on the other hand, the chi-square difference is large, then the new restriction imposed may be rejected in favor of the original model (the model with fewer constraints on its parameters).

Results and Discussion

The initial model tested in the present study (M1) postulated the equivalence of the measurement properties (equality of factor loadings, equality of the variances of measurement errors) as well as that of the errors in the structural equation, across the two groups. This initial model can be stated more formally as follows: Λ_y, Θ_ε, Ψ and ν_i are invariant across the two groups.

Of the 27 parameters to be estimated, 18 are invariant across the two groups: 3 factor loadings, 6 uniquenesses, 3 regression residuals and 6 origins of measurements. The remaining 9 are: 6 betas (3 for each group) and 3 alphas (for one group only) are to be estimated. Since there were 54 unique elements in the input data (each group had 15 correlations, 6 standard deviations and 6 means) and 27 parameters to be estimated, the model had 27 degrees of freedom (54 - 27 = 27). The chi square for M1 is 33.64 with 27 degrees of freedom, indicating that the fit is acceptable.

A more restricted model (M2) was then tested by fixing the parameter β_{64} to zero for the Regular group. M2 is in effect a quasi-simplex model for the Regular group. This new model yielded a chi square of 34.92 with 28 degrees of freedom. The difference between the chi-square values for M1 and M2 is not significant, indicating that the additional restriction imposed is reasonable, i.e., that a quasi-simplex model is indeed adequate for the Regular group.

To determine whether the quasi-simplex model was adequate for both groups, the next model tested (M3) was one in which the beta parameter β_{64} was set to zero for the Immersion group as well. This model is also acceptable, as indicated by both the chi square and the chi-square difference (M3-M2).

A restriction was then imposed on M3 by constraining the parameter β_{54} to be invariant across the two groups. The resulting model (M4) yielded a chi square of 34.93 with 30 degrees of freedom. The difference between the chi-square values for M4 and M3 with 1 degree of freedom is not significant, indicating that the hypothesis that the two groups have equal β_{54} parameters cannot be rejected. This establishes that the relationships between Grade 5 and Grade 4 are the same in both groups.

The next model tested (M5) is a test of the hypothesis that the relationship between Grade 5 and Grade 6 are the same for the two groups. This model is consistent with the data (χ^2 = 34.94, d.f. = 31). Furthermore, the difference between the chi-square values for the last two models (M4 and M5) is only 0.01. Consequently, the two non-zero beta coefficients, β_{54} and β_{65}, can be regarded as equal over the two groups. The conclusion is that the entire structural relationship (except the intercepts) of the true variables is identical for both groups. (See Table 3.)

Table 3

Chi Square Difference Tests

Model	χ^2	d.f.	p	χ^2 d.f.=1	p
1. M1:	33.64	27	.177	-	-
2. M2: M1 and $\beta^R_{64}=0$	34.92	28	.172	1.28	n.s.
3. M3: M2 and $\beta^I_{64}=0$	34.93	29	.207	0.01	n.s.
4. M4: M3 and $\beta^R_{54}=\beta^I_{54}$	34.93	30	.243	0.00	n.s.
5. M5: M4 and $\beta^R_{65}=\beta^I_{65}$	34.94	31	.286	0.01	n.s.
6. M6: M5 and $\alpha_5=0$	42.11	32	.109	7.17	.01
7. M7: M5 and $\alpha_6=0$	38.23	32	.208	3.29	.10

Once it has been shown that the betas are equal for the two groups, it is meaningful to consider the parameter alpha as an estimate of the difference between the two programs.

A further model (M6) was then constructed to test the hypothesis that the difference between the two programs is zero at Grade 5. This model yielded a chi square of 42.11 with 32 degrees of freedom, a substantially worse fit than M5. The difference between the chi-square values for M6 and M5 is 7.17 with 1 degree of freedom, which is significant at the .01 level. This would indicate that the difference between the two programs at Grade 5 (α_5 = 2.518) is statistically significant.

The next model constructed (M7) tests the hypothesis that the difference between the two programs at Grade 6 (alpha6) is zero when the initial differences are controlled for. This model yielded a chi square of 38.23, with 32 degrees of freedom. However, the difference between the chi-square values for M5 and M7 is 3.29 with 1 degree of freedom (p = .10) which would indicate that the difference between the two programs at Grade 6 (α_6 = 2.15) is marginally significant.

To summarize, the analysis shows that the measurement properties of the tests were the same for the two groups. In both groups the relationships between Grades 4 and 5 and between Grades 5 and 6 from the aspect of mathematics ability can be described by the same quasi-simplex model. The results also indicate that the two groups started with equivalent mathematics ability. As shown in Table 4 the Immersion group had a significant higher level (2.52 points) of growth in mathematics ability than the Regular group at Grade 5; in the subsequent year, however, the Regular group exceeded the Immersion group by 2.15 points in its average growth. In comparison with the Regular group, the Immersion group grew more rapidly from Grade 4 to 5 and more slowly from Grade 5 to 6.

Table 4

Maximum Likelihood Estimates for M5

ν_5	72.969 (0.970)
ν_6	67.955 (0.908)
ν_3	61.387 (0.788)
ν_4	58.523 (0.849)
ν_1	49.380 (0.757)
ν_2	47.821 (0.766)
λ_6	0.846 (0.064)
λ_4	1.036 (0.074)
λ_2	0.965 (0.069)
β_{54}	1.202 (0.084)
β_{65}	0.929 (0.073)
ζ_6	13.526 (5.676)
ζ_5	10.460 (3.344)
σ^2_ξ	64.869 (8.674)

	Regular	Immersion
α_6	2.146 (1.166)	0
α_5	-2.518 (0.930)	0
κ_4	-0.583 (1.364)	0

$\chi^2 = 34.94$ d.f. = 31 p=0.286

These results were arrived at by testing several models sequentially, each model being more restrictive than its predecessor. This procedure is essentially an exploratory one. To confirm the results obtained in this study it would be necessary to test the model on a different set of data.

References

Bentler, P.M. Multivariate analysis with latent variables: Causal modeling. In Fornell, C. (Ed.) A second generation of multivariate analysis. New York: Praeger, 1982.

Genessee, F. Immersion experiments in review. Applied Psycholinguistics, 4, 1-46, 1983.

Jöreskog, K.G. Simultaneous factor analysis in several populations. Psychometrika, 36, 239-251, 1971.

Jöreskog, K.G. Statistical models and methods for analysis of longitudinal data. In Magidson, J. (Ed.) Advances in factor analysis and structural equation models. Cambridge, Mass.: Abt Books, 1979.

Jöreskog, K.G., & Sörbom, D. LISREL V—Analysis of linear structural relationships by the method of maximum likelihood. Chicago: National Educational Resources, 1981.

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

Sörbom, D. An alternative to the methodology for analysis of covariance. Psychometrika, 43, 381-396, 1978.

Sörbom, D. Structural equation models with structured means. In Systems under indirect observation: Causality-Structure-Prediction. Netherlands: North Holland Publishing, 1982.

Swain, M., & Lapkin, S. Bilingual education in Ontario: A decade of research. Toronto: The Ministry of Education, Ontario, 1981.