J.J. Hox (1993). Factor analysis of multilevel data: Gauging the Muthén model. In: J.H.L. Oud & R.A.W.van Blokland-Vogelesang (Eds.), *Advances in longitudinal and multivariate analysis in the behavioral sciences*, chapter 10 (pp.141-156). Nijmegen, NL: ITS.

FACTOR ANALYSIS OF MULTILEVEL DATA. GAUGING THE MUTHÉN MODEL.

Joop J. Hox*
University of Amsterdam

Abstract

Social data often have a hierarchical structure. A familiar example is educational research data, with their distinct pupil, class and school levels. In the most general case there are variables defined at each level, and the variable set may be different at the different levels.

Even if all variables are measured at the lowest level, the clustering in the data leads to problems, because the usual assumptions of independently and identically distributed variables are not met. Muthén has described an approximate model for factor analysis and path analysis with hierarchical data. This chapter presents a model for factor analysis of hierarchical data proposed by Härnqvist and the model proposed by Muthén model. The accuracy of the approximation of the Muthén model is assessed by gauging.

1. Introduction

Social science often studies systems that possess a hierarchical structure. Naturally, such systems can be observed at different hierarchical levels. Familiar examples are the educational system, with its hierarchy of pupils within classes within schools, families, with family members within families, and other social structures where individuals are grouped in larger organizational or geographical groups. As a consequence the data can be regarded as a multistage or cluster sample from different hierarchical levels. In the most general case, there are not only variables at separate levels, but there may be different sets of variables at the separate levels.

**I thank Edith de Leeuw, Bengt Muthén, and two anonymous reviewers for their comments on earlier versions, Arie van Peet for his permission to use his data, and Jost Reinecke for helpful comments on the Lisrel-implementation. Special thanks are due to Bengt Muthén for his permission to participate in one of his stimulating graduate seminars at UCLA.*

Even if the analysis includes only variables at the lowest (individual) level, standard multivariate models are not appropriate here. The hierarchical structure of the population, which is reflected in the sample data, creates problems, because the standard assumption of independent and identically distributed observations (i.i.d.) is most probably not true. Multilevel analysis techniques have already been developed for the hierarchical linear regression model, and specialized software is now widely available (cf. Mason, Wong & Entwisle, 1984; De Leeuw & Kreft, 1986; Raudenbush & Bryk, 1986; Longford, 1987; Goldstein, 1987). The more general case of multilevel analysis of covariance structure models has been discussed by, among others, Goldstein and McDonald (1988), Muthén and Satorra (1989), and Muthén (1989, 1990). The approach by Muthén (1990) is particularly interesting, because many models can be analyzed with available covariance structure analysis software (such as Lisrel, Liscomp, EQS).

This chapter concentrates on the two-level factor analysis model, which assumes that we have a number of variables, measured at the lowest (individual) level, and want to determine and/or compare the factor structure at both levels. The first model discussed is a decomposition model proposed by Härnqvist (1978). The next model is the confirmative factor analysis model developed by Muthén. Subsequently, the accuracy of the Muthén model is gauged by applying it to a two-level data set with a known multilevel factor structure. The discussion provides some suggestions for the analysis of more general models.

2. The Härnqvist Model

Härnqvist (1978) proposes to use Cronbach and Webb's (1975) decomposition of the observed total scores at the individual level Y_T into a between group component Y_B, which equal the disaggregated group means, and a within group component Y_W, which equal the individual deviations from the corresponding group means. This leads to additive and orthogonal scores for the two levels (cf. Cronbach & Webb, 1975). Thus, at the individual score level we have

$$Y_T = Y_B + Y_W \quad (1),$$

while for the sample covariance matrix S(Y) we have

$$\mathbf{S}_T = \mathbf{S}_B + \mathbf{S}_W \quad (2).$$

Härnqvist recommends to scale both matrices $\mathbf{S}_B$ and $\mathbf{S}_W$ by dividing each element by the product of the standard deviations of the total scores. This results in two covariance matrices R_B and R_W, which are orthogonal and sum to the sample correlation matrix for the total scores R_T. The matrices R_B and R_W are analyzed with standard component or explorative factor analysis techniques. (As an added refinement, Härnqvist recommends to use the proportion of variance at each level as the communality estimate.)

Härnqvist's decomposition model provides a straightforward technique to explore the sample factor structure at two (or more) levels. In addition, it is easy to run with well-known software packages such as SPSS or SAS. It's main disadvantage is that it is purely explorative; it does not address problems of statistical estimation and inference. Muthén's approach, which is discussed in the next section, is more general, because it is designed to model multilevel population structures by maximum likelihood estimation.

For applications of the Härnqvist decomposition model, see Härnqvist (1978) and Hox & Willemse (1985).

3. The Muthén Model

In Muthén's model specification, we assume sampling at two levels, with both between group (group level) and within group (individual level) covariation. Thus, in the population we can distinguish the between group covariance matrix Σ_B and the within group covariance matrix Σ_W. Muthén (1989, 1990) formulates between and within structural equation models for Σ_B and Σ_W, and derives maximum likelihood procedures to fit these. The general likelihood equation is very complicated, but the likelihood for the confirmative factor analysis model turns out to be comparatively conventional (Muthén, 1990, 1991), which allows us to analyze the confirmative factor model using conventional software.

In the special case of G balanced groups, with all G group sizes equal to n, and total sample size N = Gxn, we can define two sample covariance matrices: the pooled within covariance matrix $\mathbf{S}_{PW}$ and the between covariance matrix $\mathbf{S}_B$, which are given by:

$$\mathbf{S}_{PW} = \frac{\sum^{G} \sum^{n} (Y_{gi} - \bar{Y}_g)(Y_{gi} - \bar{Y}_g)'}{N - G} \qquad (3)$$

and

$$\mathbf{S}_B = \frac{n \sum^{G} (\bar{Y} - \bar{Y}_g)(\bar{Y} - \bar{Y}_g)'}{G} \qquad (4).$$

In 3 and 4 $\mathbf{S}_{PW}$ is the covariance matrix of the deviation scores, with denominator N-G instead of N-1, and $\mathbf{S}_B$ is *n* times the covariance matrix of the group means, with G instead of the more regular G-1 as the denominator. As Muthén (1989, 1990) shows, in the balanced case $\mathbf{S}_{PW}$ is the maximum likelihood estimator of Σ_W, with sample size N-G, while $\mathbf{S}_B$ is the maximum likelihood estimator of the composite $\Sigma_W + c\Sigma_B$, with sample size G and c equal to the common group size n:

$$\mathbf{S}_{PW} = \Sigma_W \qquad (5)$$

and

$$\mathbf{S}_B = \Sigma_W + c\Sigma_B \qquad (6).$$

Equations 3 through 6 suggest using the multi-group option of conventional covariance structure analysis software to carry out a simultaneous confirmative factor analysis at both levels.

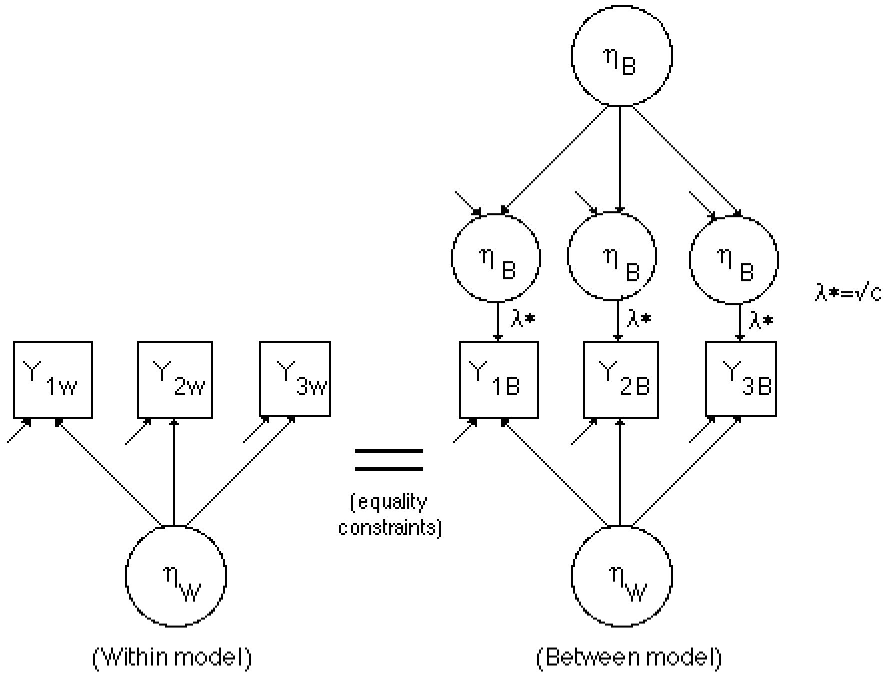

Figure 1. Single factor Within and Between Model

For the within group structure, this requires that the same model is specified for $\mathbf{S}_{PW}$ and $\mathbf{S}_B$, with equality constraints across both groups. For the between group structure, a model must be specified which incorporates the constant c as a scaling factor. Figure 1 below presents a Lisrel model specification for three observed variables and a one-factor model for both the within group and between group structure. Note that the 'between group structure' is actually a composite of the model for Σ_W and the model for Σ_B, with a scaling parameter for the latter.

The unbalanced case, with G groups of unequal sizes, is more complicated. In the unbalanced case, $\mathbf{S}_{PW}$ can be shown to be the ML estimator of Σ_W:

$$\mathbf{S}_{PW} = \Sigma_W \qquad (7)$$

but $\mathbf{S}_B$ now estimates a different expression for each separate group size *d*:

$$S_{Bd} = \Sigma_W + c_d\Sigma_B \qquad (8),$$

where 8 holds for each set of groups with a distinct common group size equal to n_d, and $c_d=n_d$ (Muthén, 1990, 1992).

The Full Information Maximum Likelihood (FIML) solution for 8 is to specify as many between-group models as there are distinct group sizes, with different scaling parameters c_d and equality constraints for the other parameters in the Between model.** This results in large and complex covariance structure models. As a simplification, Muthén (1990, 1992) proposes to utilize a partial ML approach (Muthén's approximate ML solution, or MUML for short), by computing one singe $\mathbf{S}_B$ following equation 4, and use an ad hoc estimator C* for the scaling parameter c (usually C* is close to the mean group size):

$$C^* = \frac{N^2 - \sum^{G} n^2_g}{N\,(G-1)} \qquad (9).$$

For confirmative factor analysis models, the FIML solution is exact, and MUML is an approximation which should be reasonable if the distinct group sizes n_g are not too dissimilar. Muthén (1990) presents some examples where the different approaches result in practically identical parameter estimates and chi-square values. In the next section, a different approach is used to gauge the accuracy of the MUML solution.

The decomposition model proposed by Härnqvist has some similarity to the Muthén model. If the group sizes are equal, the estimators $\mathbf{S}_{W\,(H)}$ and $\mathbf{S}_{B\,(H)}$ used by Härnqvist have a simple relationship to the estimators $\mathbf{S}_{W\,(M)}$ and $\mathbf{S}_{B\,(M)}$ used by Muthén:

$$\mathbf{S}_{W(H)} = \frac{N-G}{N}\,\mathbf{S}_{PW(M)} = \frac{N-G}{N}\,\Sigma_W \qquad (10),$$

and

$$\mathbf{S}_{B(H)} = \frac{G}{n}\,\mathbf{S}_{B(M)} = \frac{G}{n}\,\Sigma_W + \Sigma_B \qquad (11).$$

Thus, the Härnqvist estimator for $\mathbf{S}_W$ is off by a scale factor. If the number of groups G is small relative to the number of individual observations N, the approximation is good. If the correlation matrix is analyzed instead of the covariance matrix, the Härnqvist estimator is equal to the Muthén

***. This model could of course be generalized to model different between- group factor structures for groups of different sizes, but in my view this should only be attempted if there is a theoretical justification for this.*

estimator. Analyzing a scaled version of $S_{W\,(H)}$ by explorative component analysis appears a useful procedure for exploration.

The Härnqvist estimator for $\mathbf{S}_B$ also differs by a scale factor from the Muthén estimator. As Muthén (1989) has shown, such estimators confound the within group and between group covariance. Again, if the number of groups G is small compared to the number of observations N, then the number of observations n in each group is large, and $S_{B\,(H)}$ will be a reasonable approximation to $\mathbf{\Sigma}_B$, especially if the between group covariances are relatively large. All in all, the Härnqvist approach appears useful for exploration, especially if the between group covariances are large and the average group size is not small.

4. Gauging and the Gauge-model

Gauging is a process to probe the merits of a specific technique, by constructing a model with known properties, applying the technique, and studying how well the technique recovers the known properties (cf. Gifi, 1990, p34). To gauge the MUML solution, data were generated from a known factor model, the Nine Psychological Variables example from the Lisrel 7 manual (Jöreskog & Sörbom, 1989, p104). The population model is given in Table 1 below:

Table 1
Known Values Nine Variables Factor Model

Var:	Factor Matrix (LX):			Unique Var. (TD)
1	.708			.498
2	.483			.767
3	.649			.578
4		.868		.247
5		.830		.311
6		.825		.319
7			.675	.545
8			.867	.248
9	.459		.412	.471

Factor Correlations (PH)

1.000		
.558	1.000	
.392	.219	1.000

Several different two-level factor models were derived from the known factor model in Table 1. In all cases, the parameters of the within group model were specified to be equal to the values in table 1. The between group model was specified to be equal to the within model, and then scaled to produce three different two-level models, with population intraclass correlations (rho) of the observed variables equal to .25, .50, and .75. In other words: these three intraclass correlations specify population models in which the variance of the latent variables for the between group model are 30%, 100%, and 300% of the variances of the corresponding latent variables in the within group model. From the resulting three known factor models, three population correlation matrices were derived for the nine observed variables. The raw scores were generated using the normal distribution. Four different sampling schemes were used with varying group sizes: a group size ranging from 3-9, that is intended to reflect applications such as family research, and a group size ranging from 20-30, that is intended to reflect applications such as school research. The group sizes were (approximately) uniformly distributed over the group size range. The four sampling schemes are summarized in Table below 2:

Table 2
Sampling Schemes Used to Generate Data

N	n_g	Size range
300	50	3-9
600	100	3-9
1250	50	20-30
2500	100	20-30

Combining the three population models characterized by intraclass correlations of .25, .50, and .70, with the four sampling schemes in Table 2, produces 12 different data sets.

4.1 Recovery of Parameter Values by Muthén's partial ML solution

Figure 2 presents the two-factor model in Lisrel-notation in a unified form:

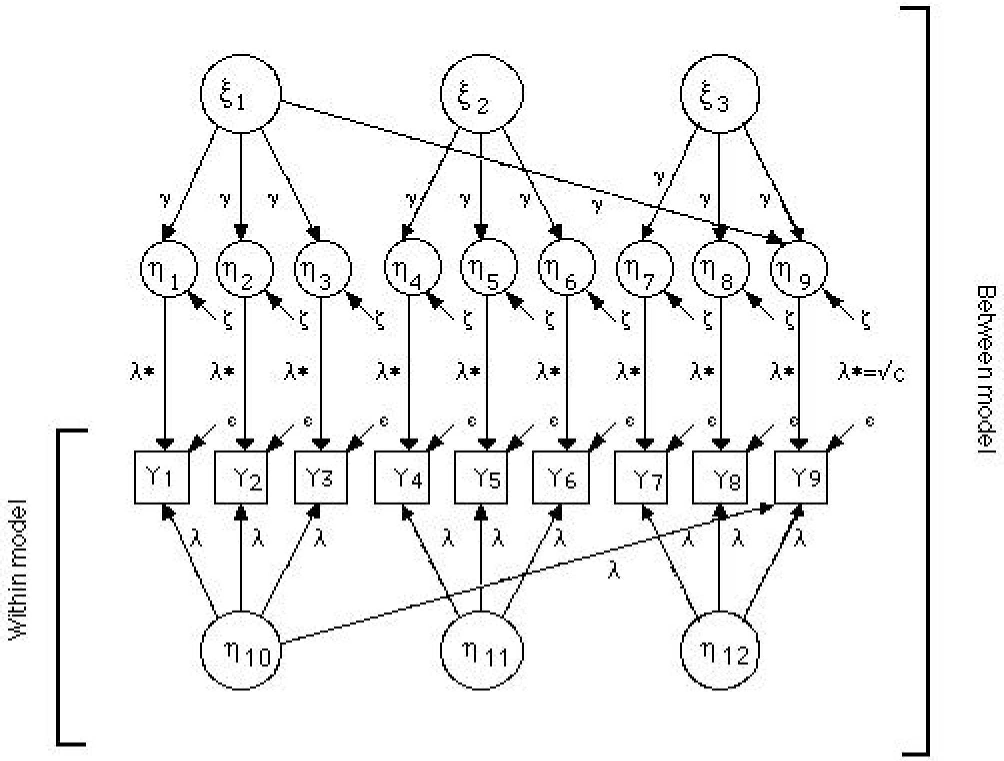

Figure 2. Gauge Model (Within and Between Group Model)

In Figure 2, there are equality restrictions for the within group model across the two levels. The between part of the model is fixed to zero in the within group. The scales of the latent variables are fixed by standardization to a variance equal to one. Maximum likelihood estimation is used to estimate the parameter values. To facilitate the comparison of solutions with different scalings for the between-model, all between group parameter estimates have been rescaled to the scale of the original model in Table 1.

Table 3 below presents the results for the within group factor structure. For all analyses, Table 3 presents the mean value of the parameter bias (mean deviation of the parameters from their true values), and the Root Mean Square Deviation RMSD (root mean squared deviation of the parameters from their true values):

Table 3
Recovery of Within Structure.

	Rho = .25		Rho = .50		Rho = .75	
	bias	RMSD	bias	RMSD	bias	RMSD
LX 300/50	.01	.11	.01	.11	.01	.11
600/100	-.01	.05	-.02	.05	-.02	.05
1250/50	-.02	.03	-.02	.03	-.02	.03
2500/100	-.00	.02	-.00	.02	-.00	.00
PH 300/50	.01	.01	.01	.01	.01	.01
600/100	-.03	.03	-.02	.03	-.02	.03
1250/50	-.03	.03	-.03	.04	-.03	.04
2500/100	-.03	.03	-.03	.03	-.03	.03
TD 300/50	-.03	.05	-.03	.05	-.03	.05
600/100	-.02	.03	-.02	.05	-.02	.05
1250/50	-.01	.02	-.01	.02	-.01	.02
2500/100	-.00	.01	-.01	.01	-.00	.01

Table 3 shows that for the within group structure, the mean bias is in all cases rather small. As the RMSD values show, the deviation of a particular parameter can still be considerable. This is especially true for the factor loadings when the within group sample size is comparatively small. With large sample sizes (600 and up), the deviation of individual loadings is also rather small.

Table 4 below presents the results for the between group factor structure. For all analyses, Table 4 presents the mean value of the parameter bias (mean deviation of the parameters from their true values), and the Root Mean Square Deviation RMSD (root mean squared deviation of the parameters from their true values):

Table 4
Recovery of Between Structure
(rescaled to same scale as the original structure)

	Rho = .25		Rho = .50		Rho = .75	
	bias	RMSD	bias	RMSD	bias	RMSD
LX 300/50	.02	.19	.01	.11	.01	.06
600/100	-.02	.08	-.02	.05	-.01	.03
1250/50	-.03	.04	-.02	.03	-.01	.01
2500/100	-.00	.04	-.00	.02	-.00	.01
PH 300/50	.01	.01	.01	.01	.01	.01
600/100	-.03	.03	-.02	.03	-.02	.03
1250/50	-.03	.04	-.03	.04	-.03	.04
2500/100	-.03	.03	-.03	.03	-.03	.03
TD 300/50	-.05	.08	-.03	.05	-.02	.03
600/100	-.04	.06	-.02	.05	-.01	.03
1250/50	-.01	.03	-.01	.02	-.01	.01
2500/100	-.01	.02	-.01	.01	-.00	.01

Table 4 shows that the recovery of the between group structure is not as good as the recovery of the within group structure, especially when the sample size is small and the intraclass correlation is low. The RMSD for the between structure is generally larger than for the within structure. A comparison of the different sampling schemes shows that the problem lies not exclusively with the lower effective between group sample size that results when groups are the units of observation. Both for 50 and for 100 groups, the between group results become much more stable when the within group sample sizes become larger. The explanation is that the between group model effectively models the covariances that are not explained by the common within group model. Table 4 reveals the importance of obtaining a correct within groups model for the accuracy of the between group parameter estimates.

Since the goal of the analysis is to separate the between group and within group effects, it is interesting to see how well this procedure recovers the correct value of the intraclass correlation, which is the population proportion of the between group variance. In the unbalanced Muthén model, this is for each variable estimated by the ratio:

$$ICC = \frac{(s^2_B - s^2_W) / c^*}{(s^2_B - s^2_W) / c^* + s^2_W} \quad (12)$$

Table 5 below shows how well Muthén's approach recovers the known population intraclass correlation (ICC):

Table 5
Recovery of ICC by Muthén's Approach

	Rho = .25		Rho = .50		Rho = .75	
	bias	RMSD	bias	RMSD	bias	RMSD
LX 300/50	.01	.06	.01	.06	.01	.04
600/100	.03	.04	.03	.04	.02	.03
1250/50	.01	.04	.02	.06	.02	.04
2500/100	.03	.04	.03	.02	.02	.03

While the estimates of the intraclass correlations are quite accurate, Table 5 suggests that they are consistently too large, and that the size of the bias depends mostly on the number of groups. It is interesting to note that if the more familiar procedure is followed to estimate the intraclass correlations by a oneway analysis of variance, the results are very similar:

Table 6
Recovery of ICC by Oneway Analysis of Variance

	Rho = .25		Rho = .50		Rho = .75	
	bias	RMSD	bias	RMSD	bias	RMSD
LX 300/50	.01	.06	.02	.06	.02	.04
600/100	.03	.04	.03	.05	.03	.04
1250/50	.02	.04	.03	.06	.03	.05
2500/100	.03	.04	.04	.05	.03	.03

Oneway analysis of variance estimates the ICC with slightly less accuracy than the more intricate Muthén approach.

As figure 1 and 2 show, the c* is basically a scaling factor, and its accuracy is only very important if the goal of the analysis is to compare different between group structures, e.g., for small versus large groups.

5. An Empirical Example

The data of this example analysis stem from the dissertation study of Van Peet (1992). The example data are the scores of 187 children from 37 large families (an average of 5 children in each family) on 6 subtests of the Groninger Intelligence Test (GIT). The subtests are: wordlist, laying cards, matrices, hidden figures, naming animals, and naming occupations. The data form a hierarchical structure, with children nested within families. Since intelligence generally shows strong effects of shared hereditary and environmental factors, strong family effects are to be expected. The scores on the subtests have been decomposed into group level and individual level variables after Cronbach and Webb (1975) (cf. equations 1 and 2). The mean and variances of the subtests at the separate levels are given in Table 7 below.

Table 7
Mean, Variances at Separate Levels,
and Intra Class Correlations (ICC) for Family Data

	Total		Family	Indiv.	
Subtest	Mean	Var.	Var.	Var.	ICC
word list	29.80	15.21	7.48	7.73	.37
cards	32.68	28.47	13.65	14.82	.35
matrices	31.73	16.38	5.24	11.14	.15
hidden figs.	27.11	21.23	6.84	14.38	.16
list animals	28.65	22.82	8.46	14.36	.22
list occup.	28.28	21.42	9.11	12.31	.28

It is clear that there is considerable family level variance. To analyze the factor structure of the six subtests a within family covariance matrix $\mathbf{S}_W$ and a between family covariance matrix $\mathbf{S}_B$ were computed following Muthén's approach given in equations 3 and 4.

The first step in the analysis is to model the within family covariance matrix. To obtain some information on the factor structure of the within family model, a component analysis was performed on the correlation matrix of the individual deviation scores (this is equivalent to Härnqvist's approach, and can easily be performed using a standard statistics package such as SPSS). The exploratory analysis a two-factor structure for the within family matrix, with the first three variables loading on the first factor, and the last three variables loading on the second factor. To check this exploratory analysis, a confirmatory factor analysis was performed on the within family covariance matrix. A model with all variables loading on one general factor was rejected (chi-squared=44.87, df=9, p=.00), while a two factor model with the first three variables loading the first factor and the last three variables loading the second factor was accepted (chi-squared=7.21, df=8, p=.51).

The two-factor model was used as the starting point for the multilevel factor analysis. As a first step, two models were analyzed using the multigroup specification proposed by Muthén. In both models, the two-factor model found above is specified for both the within family covariance matrix $\mathbf{S}_W$ and the between family covariance matrix $\mathbf{S}_B$, with appropriate equality restrictions. In addition, the first model estimates in the between family matrix $\mathbf{S}_B$ an unrestricted between family model, by estimating all covariances between the between family factors that represent the between family parts of the observed variables (in other words: no restrictions are placed on that part of the matrix Psi). This is the *maximal model*: it places no restrictions on the between family structure, and estimates the within model using the information in both the within family covariance matrix $\mathbf{S}_W$ and the between family covariance matrix $\mathbf{S}_B$. The second model is the *minimal model*: it contains *no* between family model, and again estimates the within model using the information in both the within family covariance matrix $\mathbf{S}_W$ and the between family covariance matrix $\mathbf{S}_B$. The minimal model, in fact, assumes that in the population the between family covariance matrix Σ_B is zero. Since the within part of the model holds in an analysis of $\mathbf{S}_W$ only, it is to be expected that the maximal model will be accepted in the multigroup analysis as well. And, since Table 7 shows the amount of between family variation to be large, it is to be expected that the minimal model will be rejected. As a first approach to modeling the between structure two models are examined that lie between the extremes of the minimal and the maximal model: a first that specifies one single general factor for the between family model, and a second that specifies a two factor structure similar to the two factors in the within model.

Table 8 below gives the results for the minimal and the maximal model, together with the results of the one- and two-factor model:

Table 8
Comparison of Three Between Family Models

Between family model	Chi-squared	df	p
Minimal model	125.41	29	.00
One factor model	21.28	17	.21
Two factor model	20.06	16	.22
Maximal model	7.21	8	.51

The minimal model, which specifies no between model, is rejected. The maximal model is accepted; its parameter estimates for the within factor structure are very close to the estimates obtained by analyzing only the within covariance matrix.

The one-factor model has a satisfactory fit. The fit of the two-factor model is not significantly better than the fit of the one-factor model. It seems that a one factor configuration for the between part of the model is sufficient.

Table 9 below shows the parameter estimates for the one-factor model. For the purpose of interpretation, all parameters have been standardized to a common metric for both the within and the between part of the model.

Table 9
Within and Between Model, Standardized Factor Loadings

	Within		Between
word list	.30*		.84*
cards	.52		.78
matrices	.70		1.02
hidden figs.		.30	.58
list animals		.70	$.86^{ns}$
list occup.		.48*	.33

Correlation between the two within factors: 0.22^{ns}

* = fixed parameter; ns = not significant

Table 9 suggests that at the family level, that shows the effect of shared hereditary and environmental influences, a single general (*g*) factor explains the covariances. At the individual level, that shows the effect of idiosyncratic influences, the hypothesis of a single factor is clearly rejected in favor of a differentiation into two different factors.

6. Discussion

The results from the gauging analysis suggest that the partial maximum likelihood (MUML) multilevel confirmatory factor analysis model proposed by Muthén performs very well. When the effective within group sample size is above the lower limit (at least 200) suggested by Boomsma (1983), even with small effective sample sizes at the group level (50 and 100 in our gauging example), the results may turn out perfectly acceptable. The empirical example, which especially for the between family sample of 37 falls far below this limit, also gives results that appear quite acceptable.

The Muthén model, whether the Full Information Maximum Likelihood (FIML) or the simplified (MUML) model, is considerably more complex than the Härnqvist model. It is complex to set up, and it does not belong to the set of models for which programs like Lisrel can calculate starting values, so these must be supplied to the program. Experience has shown that one needs good starting values to allow the program to start the Maximum Likelihood iterations. In the example given above, it was pointed out that the between group model essentially models the covariances that are not explained by the common within group model. Thus, to obtain good estimates for the between group model, it is important to have a correct within group model. In the gauge example, the structure of the within group model is known. In a real world analysis, this is not the case. Muthén (1990) advises to start the model search by analyzing the total (raw) score covariance matrix $\mathbf{S}_T$ or, preferably, the pooled within covariance matrix $\mathbf{S}_{PW}$, for a first approximation. However, if the between group covariances are large, $\mathbf{S}_T$ will be quite confounded. Given the good recovery of the within group structure in Table 3, it appears more practical to start with an analysis of the pooled within covariance matrix $\mathbf{S}_{PW}$, either in a separate explorative analysis, as was done in the empirical example given above, or in a two group analysis together with the between group covariance matrix $\mathbf{S}_B$, using an unrestricted between group model, which has the advantage of using all the available information about the within group structure.

After a satisfactory within group model has been found, one can search for a between group model using a variety of approaches. Muthén (1989, 1990) gives some guidelines, using the framework of covariance structure analysis. Since the between group covariance matrix $\mathbf{S}_B$ confounds Σ_B and Σ_W, an explorative factor analysis directly on $\mathbf{S}_B$ is not an attractive option. However, it is possible to model only a within structure in a two group model, and apply explorative factor analysis to the estimated covariance matrix for the between group variables. Another possibility is to use equations 5 and 6 to correct the between group covariance matrix:

$$\mathbf{S}_B^* = \frac{\mathbf{S}_B - \mathbf{S}_{PW}}{c} \qquad (13)$$

However, $\mathbf{S}_B^*$ need not be positive definite, and both approaches appear to lead to unstable estimates. Still, as a pragmatic approach to identifying the between group structure, an explorative analysis of either estimate of the between group matrix may be appropriate.

The Muthén model, including the MUML simplification, can be generalized to include several types of linear structural models (see Muthén, 1989, 1990). If there are variables that are measured on the group level, but not at the individual level, the models become more complicated. Basically, the absence of the group level variables at the individual level is treated as a missing data problem (cf. Jöreskog & Sörbom, 1989, p258; Bollen, 1989, p370). However, if factor means are included in the model, the Muthén approach (both FIML and MUML) is probably less accurate (Muthén, 1989), and the conventional covariance structure analysis software may pose some technical problems (Muthén, 1990).

References

Bollen, K.A. (1989). *Structural Equations with Latent Variables.* New York: Wiley.

Boomsma, A. (1983). *On the Robustness of LISREL (Maximum Likelihood Estimation) against Small Sample Size and Nonnormality*. Amsterdam: Sociometric Research Foundation.

Cronbach, L.J. & Webb, N. (1975). Between-class and within-class effects in a reported aptitude x treatment interaction: Re- analysis of a study by G.L. Anderson. *Journal of Educational Psychology, 67,* 717-724.

DeLeeuw, J. & Kreft, Ita G.G, (1986). Random coefficient models. *Journal of Educational Statistics*, 11, 1, 55-85.

Gifi, A. (1990). *Nonlinear Multivariate Analysis.* New York: Wiley.

Härnqvist, K. (1978). Primary mental abilities at collective and individual levels. *Journal of Educational Psychology, 70,* 706- 716.

Goldstein, H. (1987). *Multilevel models in educational and social research*. London: Griffin.

Goldstein, H. & McDonald, R. (1988). A general model for the analysis of multilevel data. *Psychometrika*, 53, 455-467.

Hox, J.J. & Willemse, J. (1985). De empirische samenhang van sociometrische scores; een analyse op twee niveau's. *Tijdschrift voor Onderwijsresearch, 10*, 82-85.

Jöreskog, K.G. & Sörbom, D. (1989). *LISREL 7. A Guide to the Program and Applications.* Chicago: SPSS Inc.

Longford, N.T. (1987). A fast scoring algorithm for maximum likelihood estimation in unbalanced mixed models with nested random effects. *Biometrika*, 74, 817-827.

Mason, W.M., Wong, G.M. & Entwisle, B. (1984). Contextual analysis through the multilevel linear model. In S. Leinhard (ed.). *Sociological Methodology, 1983-84*. San Francisco: Jossey-Bass.

Muthén, B. (1989). Latent variable modeling in heterogeneous populations. *Psychometrika, 54,* 557-585.

Muthén, B. (1990). *Means and Covariance Structure Analysis of Hierarchical Data.* Los Angeles: UCLA Statistics series, #62.

Muthén, B. (1991). Multilevel factor Analysis of Class and Student Achievement Components. *Journal of Educational Measurement, 28,* 338-354.

Muthén, B. & Satorra, A. (1989). Multilevel aspects of varying parameters in structural models. In R.D. Bock (ed.) *Multilevel analysis of educational data*. San Diego: Academic Press.

Raudenbush, S.W. & Bryk, A.S. (1986). A hierarchical model for studying school effects. *Sociology of Education*, 59, 1-17.

Van Peet, A.A.J. (1992). *De Potentieeltheorie van Intelligentie.* (The Potentiality Theory of Intelligence). University of Amsterdam, dissertation.