

School of Education, Macquarie University

Introduction

Research into the culture of schools usually involves clustered data. For example, school climate data are very often obtained from randomly selected students within a random selection of schools. Although it is acknowledged that individuals within a school might have different beliefs, values, and perceptions regarding life in school, students from the same school are expected to have more commonalities in their views than students coming from different schools.

Several authors (e.g. XE "Mason, Wong and Entwistle" Mason, Wong and Entwistle, 1983; Raudenbush & Bryk, 1985; Aitkin & Longford, 1986; Goldstein, 1987; Bock, 1989; Bryk & Raudenbush, 1992) have pointed out that for research into schooling effects, the non-independence of the observations needs to be addressed in the analysis. These authors have also stressed the need to develop methods that can truly reflect the social organisational nature of schooling.

Educational statisticians have made significant advances in both the theory and the technology of analysing data that are hierarchical in nature since the time of Mason et al (1983). The set of approaches is known collectively as multilevel models (or hierarchical linear models), and has found extensive applications in many areas of social science research. Amongst these is the extension of structural equations for clustered data by Goldstein & McDonald (1988) in which a general model for bilevel structural equation model is discussed, Lee (1990), who developed the statistical theory in the context of maximum likelihood and generalised least squares, and Muthen (1989, 1991) who provided a protocol for implementing bilevel factor models with illustrations using both artificial and empirically obtained data.

In modelling multilevel data, the researcher is faced with the issue of units of measurement in the sense of Krane and McDonald (1978). In single-level situations, Krane and McDonald (1978) showed that pre- and post-multiplication of the correlation matrix by a scaling matrix, which is to be estimated, gives scale invariant models for a covariance matrix, i.e., the procedure yields an analysis of correlation structures. The theory is extended by Browne (1982) who provided a constrained optimisation method which gives correct standard errors, and McDonald, Parker and Ishizuka (in preparation), where a nonconstrained optimisation procedure based on the logical structure of the Reticular Action Model (McArdle & McDonald, 1984) is discussed.

In multilevel implementations, the researcher is faced with the decision of which level (for example, the school- and/or student-level) scaling should occur, if at all. In other words, whether the researcher should use the original unit of measurement, standardisation within the clusters, or standardisation between the clusters by estimating the scaling matrix either within or between the clusters, will require further understandings of what each of these models can offer in terms of

interpretation. Such discussions are beyond the scope of the present paper, but can be found in McDonald and Mok (in preparation), and in Mok and McDonald (in preparation (a)).

The study to be reported here is an application of the general bilevel structural equations model as developed in McDonald and Goldstein (1989) to the special case of bilevel factor analysis in the context of analysing a large set of severely unbalanced data on school culture.

Method

The current piece of research draws on methods discussed in McDonald and Goldstein (1989). The discussion on this section will be divided into the following subsections: (1) a review of the methods of bilevel factor analysis, (2) the sample, (3) the measures of school culture, and (4) the procedure.

1. Bilevel factor model

The general bilevel structural model given by McDonald and Goldstein (1989) is restated here for the special case of factor models (Spearman case). Suppose

y_{kji} is the observed value of variable j for student i from school k , ($j=1, \dots, J$; $i=1, \dots, n_k$, $k=1, \dots, K$)
 $y_{ki}' = [y_{k1i}, \dots, y_{kJi}]$ is the data vector for this student,
 $y_k' = [y_{k1}, \dots, y_{knk}]$ is the data vector for school k ,
 $x' = [y_1', \dots, y_K']$ is the data vector for all subjects.

Suppose also that the common factor of the J variables is denoted by Z , appropriately subscripted where necessary following the convention used for the y vector. Then the factor model can be represented by

$$y_{ki} = F_{ki} [z_{ki}, y_{ki}] = F_{ki}(I - A^*)^{-1} e^* \quad " k, i$$

where F_{ki} is the filter matrix = $[\ 0, \ I]$,

$$A^* = \begin{matrix} \hat{E} \ 0 & 0' & \cdot \\ \hat{I} & & \cdot \\ \hat{I} \ a & 0 & \cdot \end{matrix}$$

contains the transpose of the $J \times 1$ zero vector and the $J \times 1$ factor loading vector relating the measured variables to the factor,

and $e_{ki}^* = [z_{ki}, e_{ki}]$

contains the common factor score z_{ki} for student i in school k , and the $J \times 1$ vector of unique factors e_{ki} .

Then it can be easily shown that if SZ , ST and YT are the covariance matrices for the factor z_{ki} , y_{ki} and e_{ki} respectively, then

$$ST = F_{ki} (I - A^*)^{-1} S e^* (I - A^*)^{-1'} F_{ki}'$$

$$\begin{aligned}
 & \hat{E} \quad \hat{E} \quad \hat{E} \quad \hat{E} \\
 & \hat{E} \quad \hat{E} \quad \hat{E} \quad \hat{E} \\
 & = Fki \quad \hat{A}I - \hat{I} \quad \hat{I} \quad \hat{I} \\
 & \hat{I} \quad \hat{I} \quad \hat{I} \quad \hat{I} \\
 & \hat{E} \quad \hat{I} \quad \hat{I} \quad \hat{I} \\
 & \hat{I} \quad a \quad \hat{I} \quad \hat{I} \\
 & \hat{I} \quad a \quad \hat{I} \quad \hat{I} \\
 & = a \quad sz2 \quad a' + YT
 \end{aligned}$$

where \hat{I} is the zero matrix, a is the $J \times 1$ vector of factor loadings.

Multilevel factor models relax the usual assumption of e_{ki} being independently and identically distributed. Instead, y_{ki} is modelled as consisting of two components: the between-school component denoted by y_{Bki} , and the within-school component denoted by y_{Wki} . That is, the model is

$$y_{ki} = y_{Bki} + y_{Wki}$$

$$\begin{aligned}
 \text{where } y_{Bki} &= FBk [z_{Bki}, y_{Bki}] = FBk(I-AB^*)^{-1}e_{Bk}^* \\
 &= aBz_{Bk} + e_{Bk}
 \end{aligned}$$

$$\begin{aligned}
 y_{Wki} &= FWki [z_{Wki}, y_{Wki}] = FWki(I-AW^*)^{-1}e_{Wk}^* \\
 &= aWz_{Wk} + e_{Wki}
 \end{aligned}$$

" k, i

and applying the results in McDonald and Goldstein (1989), and using subscripts B and W to represent respectively the between-school and within-school components,

$$ST = SB + SW,$$

with

$$\begin{aligned}
 & \hat{E} \quad \hat{E} \quad \hat{E} \quad \hat{E} \\
 & \hat{E} \quad \hat{E} \quad \hat{E} \quad \hat{E} \\
 SB= & FBk \quad \hat{A}I - \hat{I} \quad \hat{I} \quad \hat{I} \\
 & \hat{I} \quad \hat{I} \quad \hat{I} \quad \hat{I} \\
 & \hat{E} \quad \hat{I} \quad \hat{I} \quad \hat{I} \\
 aB \quad \hat{I} & \hat{I} \quad aB \quad \hat{I} \\
 & \hat{I} \quad aB \quad \hat{I} \quad \hat{I} \\
 & = aB \quad sBz2 \quad aB' + \dots(1) \\
 & \hat{E} \quad \hat{E} \quad \hat{E} \quad \hat{E} \\
 & \hat{E} \quad \hat{E} \quad \hat{E} \quad \hat{E} \\
 SW= & FWki \quad \hat{A}I - \hat{I} \quad \hat{I} \quad \hat{I} \\
 & \hat{I} \quad \hat{I} \quad \hat{I} \quad \hat{I} \\
 & \hat{E} \quad \hat{I} \quad \hat{I} \quad \hat{I} \\
 aW \quad \hat{I} & \hat{I} \quad aW \quad \hat{I} \\
 & \hat{I} \quad aW \quad \hat{I} \quad \hat{I}
 \end{aligned}$$

$$\begin{aligned}
 & \hat{E} \quad \hat{E} \quad \hat{E} \quad \hat{E} \\
 & \hat{E} \quad \hat{E} \quad \hat{E} \quad \hat{E} \\
 SW= & FWki \quad \hat{A}I - \hat{I} \quad \hat{I} \quad \hat{I} \\
 & \hat{I} \quad \hat{I} \quad \hat{I} \quad \hat{I} \\
 & \hat{E} \quad \hat{I} \quad \hat{I} \quad \hat{I} \\
 aW \quad \hat{I} & \hat{I} \quad aW \quad \hat{I} \\
 & \hat{I} \quad aW \quad \hat{I} \quad \hat{I}
 \end{aligned}$$

$$Y^T = aW' sWz2 aW' + \dots (2)$$

where $Y^T = Y^B + Y^W$, and $sZ2 = sBz2 + sWz2$.

The factor covariance of students within each school k is $sBz2$ since for any students i and i^* with $i \in i^*$,

$$\begin{aligned} \text{Cov}(y_{ki}, y_{ki^*}) &= \text{Cov}(y_{Bki} + y_{Wki}, y_{Bki^*} + y_{Wki^*}) \\ &= sBz2 = aB' sBz2 aB' + \pi \text{ } \end{aligned}$$

Scale-invariant models in which the estimates will remain unchanged with changes in the unit of measurement of the variables (McArdle and McDonald, 1984) can be achieved by estimating a non-singular diagonal scaling matrix D , and using $y_{ki} = D' g_{ki}$ as a reparameterization of y_{ki} , where S_{gk} is constrained such that $\text{Diag}\{S_{gk}\} = I$ (McDonald, Parker, & Ishizuka; in preparation). For the Spearman factor models discussed in this current paper, the constraint reduces to (1) setting the common factor variance to unity, and (2) setting the unique factor variance of each variable to be equal to the difference between unity and the square of the common factor loadings. That is,

$$\begin{aligned} s_{zj}^2 &= 1 \\ \text{and } y_{Tjj} &= 1 - a_{j2}^2 \quad " j = 1, \dots, J \end{aligned}$$

where y_{Tjj} is the j th diagonal element of Y^T and a_j is the j th component of a . This result can be extended to multilevel situations.

In the broadest sense, the ratio $sBz2/(sBz2 + sWz2)$ can be used to denote the strength of the school on the factor model. It is possible to employ scaling for analysis of correlation structures at one or both levels. We can also choose to equate factor loadings across levels, or to estimate them independently, and to fix a factor variance to be unity at one or both levels. These choice give rise to several options which will be elaborated elsewhere. For the current discussion, bilevel factor models were fitted under each of the following specifications:

(A) No scaling at either level was used, factor loadings were allow to take on different values across levels, and both $sWz2$ and $sBz2$ were constrained to unity. This model tend to be more realistic than model B below: for example the leadership of the school principal may have a different impact to the school as an institution as compared to the impact on individual students; as well, at the school level, student-teacher relationship may be described as warm or friendly, but to individual students, the teacher may be respected for his/her professionalism as well as loved for providing pastoral care. Imposing the constraint of parallel structures across level excludes the researcher from such deeper understandings. On the other hand by allowing the factor loadings to vary freely across school- and student-levels, factor variances at both levels had to be fixed at unity and as a result, there is no direct way to compute school effect. It is possible however, to compute the ratio of $S_{ajB2}/(S_{ajB2} + S_{ajW2})$ as an

estimate of the relative amount of information contributed at the school level. The model is

$$\begin{aligned} SB &= aB \ aB' + YB \\ SW &= aW \ aW' + YW \end{aligned}$$

(B) No scaling at either level was applied, factor loadings were constrained equal across levels, and $sWz2$ was set to unity. The model is

$$\begin{aligned} SB &= a \ sBz2 \ a' + YB \\ SW &= a \ a' + YW \end{aligned}$$

In this model, the estimates were in the original metric and effect of the school is estimated by $sBz2/(sBz2 + 1)$.

(C) A third bilevel factor model to be fitted in this study is the reparameterization of Model B by scaling at both levels. That is,

$$yBk = DB \ gBk = DBFBk \ [zBk, gBk] = DBFBk(I-AB^*)-1eB^*$$

$$yWki = DW \ gWki = DFWki \ [zWki, gWki] = DFWki(I-AW^*)-1eW^*$$

" k,i

where $SgBk$ and $SgWki$ were both constrained such that $\text{Diag}\{SgBk\} = I$, and $\text{Diag}\{SgWki\} = I$. The between and within covariance matrices are:

$$\begin{aligned} \hat{\Sigma} &= \begin{pmatrix} \hat{\Sigma} & \hat{\Sigma} \theta & \theta' & \hat{\Sigma}^{-1} & \hat{\Sigma} \ sBz2 & \theta' & \hat{\Sigma} \\ \hat{\Sigma} \theta & \hat{\Sigma} \theta' & \hat{\Sigma}^{-1} & \hat{\Sigma}^{-1} & \hat{\Sigma} & \theta' & \hat{\Sigma} \\ SB= DBFBk & \hat{\Sigma} I & - & \hat{\Sigma} & \hat{\Sigma} & \hat{\Sigma} & \hat{\Sigma} I \\ \hat{\Sigma} & \hat{\Sigma} & \hat{\Sigma} & FBk' \ DB & \hat{\Sigma} & \hat{\Sigma} & \hat{\Sigma} \\ aB \ \hat{\Sigma} & \hat{\Sigma} & \hat{\Sigma} \ aB \ \hat{\Sigma} & \hat{\Sigma} & \hat{\Sigma} \ 0 & CB & \hat{\Sigma} & \hat{\Sigma} \\ & \hat{\Sigma} & \hat{\Sigma} & \hat{\Sigma} & \hat{\Sigma} & \hat{\Sigma} & \hat{\Sigma} & \hat{\Sigma} \end{pmatrix} \\ &= DB \ (aB \ sBz2 \ aB' + CB) \ DB \\ &= DB \ (aB \ aB' + CB) \ DB \dots \end{aligned} \tag{3}$$

$$\begin{aligned} \hat{\Sigma} &= \begin{pmatrix} \hat{\Sigma} & \hat{\Sigma} \theta & \theta' & \hat{\Sigma}^{-1} & \hat{\Sigma} \ sWz2 & \theta' & \hat{\Sigma} \\ \hat{\Sigma} \theta & \hat{\Sigma} \theta' & \hat{\Sigma}^{-1} & \hat{\Sigma}^{-1} & \hat{\Sigma} & \theta' & \hat{\Sigma} \\ SW= DFWki & \hat{\Sigma} I & - & \hat{\Sigma} & \hat{\Sigma} & \hat{\Sigma} & \hat{\Sigma} I \\ \hat{\Sigma} & \hat{\Sigma} & \hat{\Sigma} & FWki' \ DW & \hat{\Sigma} & \hat{\Sigma} & \hat{\Sigma} \\ aW \ \hat{\Sigma} & \hat{\Sigma} & \hat{\Sigma} \ aW \ \hat{\Sigma} & \hat{\Sigma} & \hat{\Sigma} \ 0 & CW & \hat{\Sigma} & \hat{\Sigma} \\ & \hat{\Sigma} & \hat{\Sigma} & \hat{\Sigma} & \hat{\Sigma} & \hat{\Sigma} & \hat{\Sigma} & \hat{\Sigma} \end{pmatrix} \\ &= DW \ (aW \ sWz2 \ aW' + CW) \ DW \end{aligned}$$

$$= DW (aW \quad aW' + CW) \quad DW \quad \dots(4)$$

Here, the a 's are standardised factor loadings. As such, the loadings can be compared across levels in the sense of comparing standardised regression coefficients.

(D) A fourth bilevel factor model to be fitted in this study is by estimating the scaling factors at both levels, fixing student level factor variance at unity, and constraining the factor loadings to be equal across levels. That is, equations (3) and (4) then become:

$$SB = DB (a \quad sz2 \quad a' + CB) \quad DB \quad \dots(5)$$

$$SW = DW (a \quad a' + CW) \quad DW \quad \dots(4)$$

This model enables the researcher to estimate the effect of the between factor variance, and also to have scale-invariant estimates.

2. The Sample

The sample for this study consisted of 5,932 Year-12 students selected from 50 Catholic schools in the 13 Dioceses of New South Wales and the Australian Capital Territory by means of proportionate stratified random sampling (Flynn, 1992). In the sample there were 3,048 boys and 2,884 girls. Amongst them, 1,549 of the boys were studying at Boys-only schools, 1,476 of the girls were studying at Girls-only schools, and the remaining 1,499 boys and 1,408 girls were from co-educational schools. The majority (88%) of these students were Australians by birth. About 60% of the fathers, and 57% of the mothers of students in the sample were also Australians. The majority of the remaining sample had parents originally coming from Europe or the Middle East.

Thirty-nine of the schools in the sample contained day students only, 10 schools had both day and boarding students, and the remaining school was a boarding school with no day students. Thirty of the schools were located within the Sydney region, and the others were outside Sydney. There were big schools as well as small schools in the sample. School size ranged from 200 to about 1,200 students, and from 20 to 85 full time teachers.

3. The Measures of School Culture

The scales used in this study to operationalise school culture were developed by Flynn (1982, 1992). The scales were constructed based on the premise that actors within the school define and share a common set of values, beliefs, standards and practices in their daily interactions. This set of central beliefs and norms provides meanings to the school community, and is being passed on from generation to generation. A set of school culture scales was constructed to tap three facets of school culture, viz, the educational aspects, the religious aspects, and the informal curriculum. These scales were:

(a) Quality of formal school curriculum (CURRIC)

- (b) Out of school curriculum (OUTSCH)
- (c) Images of the teachers in the school (TEACH)
- (d) Student morale (MORALE)
- (e) Images of the Principal (PRINCP)
- (f) Attitudes toward discipline (DISCIP)
- (g) Attitudes toward religious education (RE, 6 scales)
- (h) Quality of School Life (QSL; developed by A.C.E.R.; Ainley, Batten, & Miller; 1984)

The Quality of School Life Scale was intended to reflect the ethos or the climate of the school. It consists of 40 Likert-type items with a common stem which reads 'School is a place where ...' and a common 4-point response scale from Strongly Agree to Strongly Disagree, with Agree and Disagree as the middle categories. In this application two of the items were modified, and the responses were changed to a 5-point scale (Certainly False, False, Uncertain, True, Certainly True). The scale was found to have very strong psychometric properties (see, for example, Ainley, Batten, & Miller, 1986; Mok, 1992). However, It was found that for this sample at least, the Quality of School Life Scale was more suitable for measuring school life at the personal and experiential level rather than at the institutional level (Mok & McDonald, in preparation (b)). This scale is not included for the purposes of the present study.

On the other hand, whereas Religious Education formed a significant aspect of school life in Catholic schools, it was anticipated that for this particular sample -- with all the schools being Catholic schools and about 87% of the subjects in the sample were Catholics -- there might not be enough variance across schools to warrant analysis at the school level. This set of scales pertaining to Religious Education were not included in the current study.

A major function of the school is to teach academic knowledge and skills. The scholarship of the school is measured by the Quality of Formal Curriculum Scale (CURRIC), which is made up of 6 items (See Appendix 1). It addresses the issues of the quality of the formal curriculum within the school including the relevance of the subjects in meeting current and future needs, the range of subjects offered and the adequacy of the curriculum in intellectual development.

The Out-of-School Curriculum Scale (OUTSCH) is constructed to measure learning beyond academic concerns. This scale is made up of 4 items, and is concerned with the quality and form of outside classroom activities, and the student-staff relationships under such circumstances.

The Images of the Teachers Scale (TEACH) is designed to measure the extent to which teachers are perceived to carry out their teaching duties professionally, the degree to which teachers are seen to effectively mediate the school's spirit, and the degree to which teachers are perceived as friendly and supportive personnel within the school. It has 12 items.

Whether prevailing experience at school is a positive or a negative one is tapped by the Student Morale Scale (MORALE), which is made up of 10 items. The scale measures the commitment of the students toward school, and the degree to which the school climate is a happy, cohesive, and friendly one.

In the complexity of interactions among the actors within the school community, the principal serves the leadership role. Flynn (1982) summarised the importance of the principal elegantly:

" The hope and aspirations of the school community are embodied symbolically in the principal. He or she sets its tone, establishes its pattern of discipline and challenges the school community..."

(Flynn, 1982, pp 169)

The 5-item Images of the Principal scale (PRINCP) was designed to tap the extent to which the principal was perceived as a leader who shaped school lives.

The school is perhaps the first formal institution outside the home where the individual is exposed to the varied rules and regulations. The school prepares the youngster on the protocols of interaction with other actors in a formal organisation such as he/she will face in future employment. The amount of emphasis that a school places on propriety defines the press for conformity of student behaviour. This aspect of the school culture is measured by the Attitudes Toward Discipline Scale (DISCIP). It is made up of 4 items.

The scales used in this study (CURRIC, OUTSCH, TEACH, MORALE, PRINCP, DISCIP) all had a 5-point Likert type response scale which read, 'Certainly True, True, Uncertain, False, Certainly False'. With the exception of DISCIP, these were scored with values of 5,4,3,2,1 respectively. Items in DISCIP were scored in the reverse way. Scale values were formed by taking item means over the items which made up the scale. A mean value of 3 was taken to denote a neutral position, values above 3 were taken as indications of positive attitudes or feelings, while values below 3 were interpreted as negative attitudes or feelings.

4. Procedure

The values for each of the 6 school culture scales were formed by taking averages of the items making up each scale. The bilevel factor analysis conducted in this study adopted (with modifications) the strategies for multilevel covariance structure analysis recommended by Muthen (1991), and consisted of the following steps:

Step 1. Conventional factor analysis of ST. A single-level confirmatory factor analysis of the total sample covariance ST at the student level was conducted by means of COSAN (Fraser, 1987). This step was carried out with the understanding that because of the nested nature of the data, the analysis was incorrect, and the more the data are related within the schools, the more serious the error of the analysis would have been. This step aimed to gain insights into the model ideas. Muthen (1991) has pointed out that the model goodness of fit will be inflated.

Step 2. Estimation of between variation. This step aimed to check the legitimacy of multilevel factor analysis. For each scale, a variance components model was fitted to the data. That is, σ_{ykji} , the value of

student i from school k on scale j ,
 $(i=1, \dots, n_k; j=1, \dots, J; k=1, \dots, K)$, is modelled in terms of two components: (1) the within-school component which describes the scale value as a deviation from the school mean (as a result of individual differences of students), and (2) the between-school component which describes the school mean value as a deviation of school k from an overall mean value across all schools. Mathematically, the variance components model (Goldstein, 1987) is:

$$y_{kji} = a_{kj} + e_{kji} \quad (\text{the within-school model})$$

$$a_{kj} = a_j + u_{kj} \quad (\text{the between-school model}) \quad " i, j, k$$

where

a_{kj} is the mean scale value for scale j , school k ,
 e_{kji} is the within school random term for scale j
 a_j is the grand mean across all schools for scale j
 u_{kj} is the between school random term for scale j

with $e_{kji} \sim N(0, se_j^2)$
 $u_{kj} \sim N(0, su_j^2)$,
 $cov(e_{kji}, u_{kj}) = 0 \quad " i, j, k$

For each scale, the intra-school correlation, r , is given by

$su_j^2 / (su_j^2 + se_j^2)$. This value gives the proportion of total variance contributed by school differences (Goldstein, 1987). The variance component model for each of the scales was fitted by the ML3 computer software (Rasbash, J., Prosser, R. & Goldstein, H.; 1990), which minimises the loglikelihood function by the operation of the Iterative Generalised Least Squares algorithm (Goldstein, H.; 1989).

Step 3. Bilevel Factor Analysis. Bilevel factor analysis was carried out using the BIRAM program (McDonald, Middlehurst, Lam, & Parker; in preparation) which is written to implement the Reticular Action Model (McArdle & McDonald, 1984) for both single- and bilevel linear structural relations. A quasi-Newton minimisation algorithm is used in the computer software to minimise a likelihood-based discrepancy function. The program allows the option of conducting a "pseudo balanced run only", whereby the average school size is used in the procedures, versus the option of requesting a "pseudo balanced and unbalanced runs", whereby the actual school sizes are used, taking the results from the pseudo balanced run as initial values for the minimisation procedure. McDonald & Goldstein (1989) have shown that the pseudo balanced estimates are asymptotically unbiased. Given that the school sizes in the sample were widely varied -- school sizes ranged from 55 to 348 -- decision was made to report the results from the unbalanced runs. However, because the school sizes were far from balanced, it is anticipated that the optimisation procedure might take a long time and because of the modest intra-school correlation, the program might have difficulties in the computation of the standard errors for the estimates. This will have only minimal effect on our interpretation of the results because all models reported here show parameters significantly

different from zero as a result of the large sample size. Four bilevel factor models (specified as above) were fitted to the data using the RAM software. These models were compared with one another and also with conventional factor analysis of ST.

Results

Step 1. Conventional factor analysis of ST. Confirmatory factor analysis on a one-factor solution using COSAN (Fraser, 1987) gave a Chi-square value of 377.09 with 9 degrees of freedom. Because of the large sample size, the power of the test is high, and even trivial discrepancies will be detected. The sample-free goodness-of-fit index as proposed by McDonald (1989) for this factor solution is 0.97 which indicated that a single factor solution fits reasonably well. The largest residual in magnitude was 0.106, and the Root Mean Square value was 0.032. The factor loadings for this solution are given in Table 1.

Step 2. Estimation of between variation. Results of the multilevel variance components analysis using ML3 are provided in Table 2. Intra-school correlations for the six scales ranged from 7.7% to 17.6%, which were only moderate (Table 2). Column 2 of Table 2 gives the maximum likelihood estimates of the grand means of the scales. With the exception of DISCIP (Attitudes to Discipline) all scales have a grand mean above 3, which means that the students in general hold a positive view regarding the school. The school level variances were low but significantly different from zero.

Step 3. Bilevel Factor Analysis. The bilevel factor analysis results from RAM are presented in Table 1. All models were subjected to a "pseudo balanced followed by unbalanced" run. For Models C and D where scalings were used at both levels, the unbalanced runs failed to give standard errors. The parameters estimates at convergence for these models were provided.

Model A is a distinct-loading-no-scaling bilevel factor model. Chi-square value for this model is 287.76 with 18 degrees of freedom, and the corresponding McDonald's Index of Goodness of Fit is 0.98 indicating a very good fit. Because scaling was not used, the loadings were in the original metric of the covariance matrix and it would be difficult to interpret the loadings on their own. The Chi-square difference between this model and the conventional one-factor solution showed an improvement in fit in terms of Chi-square value of 89.33 points with a gain of 9 degrees of freedom. The difference suggests that the move to a bilevel treatment is probably appropriate. This means that when the hierarchical nature of the data was taken into consideration in the estimation, a better fitted model resulted. The distinct-loading-no-scaling bilevel factor model is preferred over the conventional factor solutions.

Factor loadings for the equal-loading-no-scaling model (Model B) again cannot be read on their own. The Chi-square value for this model is 293.89 with 23 degrees of freedom (McDonald's Index of Goodness of Fit is 0.977).

Comparison of the Chi-square values of Model B and Model A showed that the unrestricted Model A (which allowed factor loadings to vary freely across the levels) did not improve model fit significantly over the restricted Model B. Model B also provided an estimate on the effect of school on the factor structure. This was given by the ratio $0.11/(0.11+1)$ and was approximately 10%. The parsimonious restricted model (B) is preferred over the restricted model (A). Both bilevel factor models were better in terms of fit than the conventional model.

The estimated SB and SW were each scaled to a correlation matrix and (possibly) distinct factor loadings were estimated in Model C. The metric of the factors at both levels was determined by fixing factor variance at unity. In using scale-invariant estimates, the factor loadings can be directly compared with those obtained from the conventional factor analysis of ST at the student level. There was very little difference in terms of factor loadings between the bilevel factor solution and the conventional factor analysis. On the other hand, there was significant loss in terms of Chi-square Goodness of Fit statistics together with a gain of 9 degrees of freedom. The additional information provided by modelling the school effect was the factor loadings at the school level. At this level, all the loadings were very high -- analogous to a Heywood case in the single-level situations. When compared to either Models A or B, the unrestricted-scaling-at-both-level Model C was inferior.

By constraining factor loadings to be equal across levels and with scale-invariant estimates, the effect of the school in Model D could be estimated in terms of the metric of the variances at the school level. This was found to be 1.49, which was an equivalent of 69% of the total variances. This model appeared to fit less well than all of the models considered so far. With scaling and constrained equal factor loadings, Model D had similar loadings at student level to the conventional analysis.

Across all the comparisons, Model B (which is the restricted with no scaling model) appears to provide the best fit to the data amongst all the models considered in this analysis.

Conclusion

This study illustrated the use of bilevel factor analysis in research situations where a cluster sampling structure is used, and where the research question at the school level is to be answered by data from the student level. In terms of methodology, the study provides an empirically based example of the technology of bilevel factor modelling. If the researcher is mainly interested in student level results, then with moderate intra-school correlations such as those in this study, it would seem that there is not much difference between the estimated factor loadings obtained by the conventional approach on the one hand, and those estimated using bilevel factor analysis on the other. The major gain in employing bilevel factor analysis in such situations perhaps is that the Standard Errors are correctly estimated (Browne, 1982) in the scale-invariant models discussed. Nevertheless, if the interest of the researcher is mainly in school level factor structures, then bilevel factor analysis has a lot to offer. Not only that school level factor loadings

can be unbiasedly estimated, some models also provide an estimate of the proportion of total factor variance contributed by differences across schools. In this study, univariate distributions of some of the variables were found to be deviated from normality (see Appendix 2). How much of these deviations have affected the estimation, and how much of such deviations will affect bilevel factor solutions in general circumstances will require further investigations. Perhaps future research effort need also to be directed towards further explorations of the behaviours of the bilevel factor models in relation to the size of intra class correlations, sameness/difference of factor structures across levels, and random slope approaches.

As far as computer software is concerned, the extreme diversity of school sizes coupled with the modest sizes of the intra-school correlations in this sample has put the BIRAM computer package to a hard test. Its versatility however was demonstrated by putting various models through the system. One extension which might not be unreasonable, though, is to allow the factor models to take up different structures at the two levels. For example, it might not be unreasonable to model school culture at the organisational level as consisting of two common factors: the interpersonal factor and the scholastic factor, while maintaining a one-factor solution at the student individual-experiential level.

References

Ainley, J., Batten, M., & Miller, H. (1984). *Staying at High School in Victoria*. ACER Research Monograph No. 23. Hawthorn, Victoria: Australian Council for Educational Research.

Aitkin, M., & Longford, N. (1986). Statistical modelling issues in school effectiveness studies. *Journal of the Royal Statistical Society, Series A*, 149 (1), 1-43.

Bock, D. (1989). *Multilevel Analysis of Educational Data*. Academic Press.

Browne, M.W. (1982). "Covariance structures." In Hawkins, D.M. (ed.) *Topics in Applied Multivariate Analysis*. Cambridge University Press: Cambridge.

Bryk, A.S. & Raudenbush, S.W. (1992). *Hierarchical Linear Models*. Sage Publications.

Flynn, M. (1982). *The Effectiveness of Catholic Schools*. N.S.W. St. Paul Publications.

Flynn, M. (1992). *The Culture of Catholic Schools*. N.S.W.: St. Paul Publications. (In press).

Fraser, C. (1987). *COSAN User's Guide*.

Goldstein, H. (1987). *Multilevel Models in Educational and Social Research*. London: Griffin; New York: Oxford University Press.

Goldstein, H. (1989). "Restricted (unbiased) iterative generalised least squares estimation." *Biometrika*, 76, 622-623.

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

Krane, W.R. & McDonald, R.P. (1978). "Scale invariance and the factor analysis of correlation matrices." *Br. J. Math. Statist. Psychol*, 31, 218-228.

Lee, S.Y. (1990). "Multilevel analysis of structural equation models." *Biometrika*, 77, 4, 763-72.

XE "Mason, Wong and Entwistle" Mason, W.M., Wong, G.M. & Entwistle, B. (1983). "Contextual analysis through the multilevel linear model." In S. Leinhardt (Ed.), *Sociological Methodology*. 72-103.

McArdle, J.J. & McDonald, R.P. (1984). "Some algebraic properties of the Reticular Action Model for moment structures". *British Journal of Mathematical and Statistical Psychology*, 37, 234-251.

McDonald, R.P. (1989). "An Index of Goodness-of-Fit based on non centrality". *Journal of Classification*, 6:97-103.

McDonald, R.P. & Goldstein, H. (1989). "Balanced versus unbalanced designs for linear structural relations in two-level data." *British Journal of Mathematical and Statistical Psychology*, 42, 215-232.

McDonald, R.P., Middlehurst, J., Lam, P., & Parker, P. (In preparation). *User's Guide for the BIRAM package*.

McDonald, R.P. & Mok, M. (In preparation). "The behaviour of bilevel factor models as a function of properties of the data."

McDonald, R.P., Parker, P., & Ishizuka, T. (In preparation). "A scale-invariant treatment for recursive path models".

Mok, M. (1992). "Construct Validity of the Quality of School Life Scale". *Psychological Test Bulletin*. Vol 5, 47-51.

Mok, M. & McDonald, R.P. (In preparation (a)). "Alternative models for bilevel factor analysis."

Mok, M. & McDonald, R.P. (In preparation (b)). "Quality of school life: a scale to measure student experiences or school climate?"

Muthen, B.O. (1989). "Latent variable modeling in heterogenous populations." *Psychometrika*, 54, 557-585.

Muthen, B.O. (1991). "Multilevel Covariance Structure Analysis." in Hox, J. & Kreft I. (eds.). *Multilevel Modeling. A special issue of Sociological Methods & Research*.

Rasbash, J., Prosser, R., & Goldstein, H. (1990). *User's Guide for the ML3 Software for Three-level Analysis*. Institute of Education, University of London.

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

Acknowledgements

This research was partially supported by the Macquarie University MURG Grant No. 20623421BAL. I would like to thank Professor Roderick P. McDonald for valuable advice on the analysis and in reviewing the paper. I would also like to thank Br Marcellin Flynn for allowing me to use the set of data for this study.

Table 1. Maximum Likelihood Solutions for Conventional and Bilevel Factor Analysis

Scale

Convent-ional

Model A, Distinct loadings, no scaling

Model B, Equal loadings, no scaling

Model C, Distinct loadings, scaling at both levels**

Model D, Equal loadings, scaling at both levels**

Sch level loadings

Estimate

S.E.

Estimate

S.E.

Estimate

S.E.
Estimate
S.E.

CURRIC

-

.139
(.032)
.511
(.009)
.628
(.103)
.700
(.008)

OUTSCH

-

.174
(.053)
.523
(.010)
.486
(.123)
.639
(.009)

TEACH

-

.181
(.025)
.570
(.008)
.920
(.045)
.833
(.006)

MORALE

-

.214
(.029)
.592
(.008)
.923

(.046)
.810
(.006)

PRINCP

-

.181

(.054)

.534

(.010)

.487

(.121)

.663

(.008)

DISCIP

-

.198

(.051)

.290

(.011)

.566

(.114)

.356

(.012)

Sch level factor var

-

fixed at 1

.112

(.026)

fixed at 1

fixed at 1

Sch level resid var

CURRIC

-

.030

(.007)

.030

(.007)

.606

-

.511

-

OUTSCH

-

.098

(.021)

.095

(.020)

.764

-

.592

-

TEACH

-

.006

(.003)

.004

(.003)

.154

-

.307

-

MORALE

-

.008

(.004)

.011

(.004)

.149

-

.344

-

PRINCP

-

.105

(.022)

.107

(.023)

.763

-

.561

-

DISCIP

-

.084

(.019)

.097

(.021)

.680

-

.874

-

Sch level scale vec

CURRIC

-

-

-

-

-

.221

(.025)

.222

(.021)

OUTSCH

-

-

-

-

-

.358

(.038)

.367

(.035)

TEACH

-
-
-
-
-

.197

(.022)

.191

(.017)

MORALE

-
-
-
-
-

.232

(.025)

.222

(.020)

PRINCP

-
-
-
-
-

.372

(.039)

.340

(.037)

DISCIP

-
-
-
-
-

.351

(.037)

.339

(.034)

Stud level loading

CURRIC

.693

.512

(.009)

.511

(.009)

.700

(.008)

.700

(.008)

OUTSCH

.627

.523

(.010)

.523

(.010)

.640

(.009)

.639

(.009)

TEACH

.839

.571

(.008)

.570

(.008)

.832

(.006)

.833

(.006)

MORALE

.818

.591

(.008)

.592

(.008)

.809

(.006)

.810
(.006)

PRINCP

.641
.534
(.010)
.534
(.010)
.664
(.008)
.663
(.008)

DISCIP

.372

.289
(.011)
.290
(.011)
.354
(.012)
.356
(.012)

Stud level factor var

fixed at 1

fixed at 1

fixed at 1

fixed at 1

Stud level resid var

CURRIC

.520

.272

(.006)

.273

(.006)

.510

-

.511

-

OUTSCH

.607

.394

(.008)

.394

(.008)

.590

-

.592

-

TEACH

.297

.145

(.004)

.145

(.004)

.308

-

.307

-

-

MORALE

.331

.185

(.005)

.185

(.005)

.346

-

.344

-

PRINCP

.590
.362
(.008)
.362
(.008)
.559
-
.561
-

DISCIP

.861
.582
(.011)
.582
(.011)
.875
-
.874
-

Stud level scale vec

CURRIC

-
-
-
-
-
.731
(.007)
.731
(.007)

OUTSCH

-
-
-
-
-
.817

(.008)
.817

(.008)

TEACH

-
-
-
-
-

.686
(.006)
.686
(.006)

MORALE

-
-
-
-
-

.731
(.007)
.732
(.007)

PRINCP

-
-
-
-
-

.805
(.007)
.804
(.007)

DISCIP

-
-
-
-
-

.816
(.008)
.816
(.008)

Chi-sq
336.69
287.76

293.89

287.8

302.7

d.f.

9

18

23

18

24

** Note: These models failed to produce standard errors at convergence for the unbalanced run.

Table 2. Variance Components Models

Scale*

Grand Mean

Sch level variance

Stud level
variance

Intra-sch correlation

Estimate

S.E.

Estimate

S.E

Estimate
S.E
Estimate

CURRIC
3.644
(.033)
.049
(.018)
.534
(.010)
8.4%

OUTSCH
3.305
(.052)
.130
(.027)
.667
(.012)
16.3%

TEACH
3.675
(.030)
.039
(.009)
.471
(.009)
7.7%

MORALE
3.689
(.034)
.053
(.012)
.535
(.010)
9.1%

PRINCP
3.664
(.054)
.138
(.029)

.647
(.012)
17.6%

DISCIP
2.788
(.051)
.123
(.026)
.665
(.012)
15.6%

Note: * please refer to the text for meanings of the scales.

Appendix 1: Items making up the scales

No.

Item Stem

CURRIC - Quality of the Formal Curriculum

41

The curriculum of this school meets my present needs.

45

This school offers a good range of subjects in Years 11 and 12.

46

The subjects offered at this school develop the capacity for independent and critical thinking.

47

The subjects taught in this school offer useful knowledge or skills.

49

The subjects taught in the school are relevant to real life and to students' needs

50

The subjects taught here prepare students adequately for future employment

OUTSCH - Out-of-School Curriculum

42

There are opportunities for students to get to know teachers outside the classroom.

43

The out-of-school activities of the school have sufficient variety and scope.

44

There is a good sports program in the school.

53

The school places sufficient emphasis on cultural activities (music, art, drama, etc.)

TEACH - Images of the Teachers in the school

56

Most teachers are well qualified and have good teaching skills.

58

Most teachers in this school show a good deal of school spirit.

60

Most teachers know their Year 12 students as individual persons.

64

Most teachers are part of the school community.

65

Most teachers have a professional attitude towards their teaching.

66

Most teachers carry out their work with energy and pleasure.

72

Most teachers take a personal interest in me.

80

Most teachers give students sufficient encouragement.

83

If students have difficulty with their school work, most teachers take time to help them.

85

Most teachers go out of their way to help you.

86

Most teachers show that people are more important than rules.

90

Most teachers here are caring and willing to assist students who need help.

MORALE - Student Morale

55

Students here think a lot of their school.

69

Everyone has a lot of fun at this school.

70

A good spirit of community exists amongst Year 12 students.

73

I have been happy at school.

76

Students at this school do not mind wearing the school uniform.

79

Everyone tries to make you feel at home in this school.

87

My experience of this school has been a happy one.

91

I am happy to be a student at this school.

92

School rules here encourage self-discipline and responsibility.

93

There is a happy atmosphere in the school.

PRINCP - Images of the Principal.

62

I can approach the Principal at any time for advice and help.

71

The Principal ensures that the school provides a good education to students.

81

The Principal encourages a sense of community and belonging in the school.

89

The Principal places importance on the religious nature of the Catholic

school.

94

The Principal provides good leadership of the school community.

DISCIP - Attitudes toward Discipline

59*

Year 12 students are not given sufficient real freedom here.

74*

This school places too much emphasis on external conformity to rules and regulations.

78*

Most teachers never explain why they ask you to do things around here.

84*

There are too many rules which restrict students' freedom.

Note: * These items were reversely coded.

Appendix 2. Sample Descriptive Statistics for the Scales
(n=5,932)

CURRIC
OUTSCH
TEACH
MORALE
PRINCP
DISCIP

Mean

3.690

3.354

3.692

3.717

3.652

2.819

(SE Mean)

(.010)

(.012)

(.009)

(.010)
(.012)
(.012)

Stand Dev

.764
.895
.714
.767
.893
.888

Variance

.584
.800
.509
.588
.797
.789

Kurtosis

.331
-.384
.635

.489
-.052
-.544

(SE Kurt)

(.064)
(.064)
(.064)
(.064)
(.064)
(.064)

Skewness

-.579
-.331
-.642
-.739
-.615
.083

(SE Skew)

(.032)
(.032)
(.032)
(.032)
(.032)
(.032)

Minimum

1.00
1.00
1.00
1.00
1.00

1.00

Maximum

5.00

5.00

5.00

5.00

5.00

5.00

AARE '92 Multilevel Ana of Sch Culture Page PAGE27 of 27,

PAGE \# "'Page: '#'

'"Page: 2

PAGE \# "'Page: '#'

'"Page: 21

ÄÉ-Ñ-s`t`rtosis

.331

-.384

.635

.489

-.052

-.544

(SE Kurt)

(.064)

(.064)

(.064)

(.064)

(.064)

(.064)

Skewness

-.579

