Eğitim Yönetimi
Yıl 2, Sayı 3, Yaz 1996, ss.409-420

MULTILEVEL ANALYSIS OF HIERARCHICALLY STRUCTURED EDUCATIONAL DATA

Dr. Emin KARİP

The multilevel data analysis, a conceptually and technically appealing statistical procedure, is becoming increasingly popular among researehers vvho focus on evaluation of educational programs, smdent achievement and assessment of grovvth in achievement (Sheltzer, 1995). Although application of multilevel model in econometrics, biometrics. and sociological research is not novel, its application to educational data is a very recent development. To date. there is only a small but grovving number of studies utilizing Üıis model in boüı developing and developed countries (Aitkin & Zuzovsky, 1991. 1992; Bernstein, 1990: Goldstein, 1989; Lockheed & Longford. 1989. 1991: Mislevy & Bock. 1989; Raudenbush & Bryk, 1989, 1992: Riddel, 1989).

Since the fırst publication of Coleman Report in 1966, Üıere has been an increasing research focus on examining the determinants of student aclıievement. Research fındings regarding sources of variation in smdent achievement among schools have been inconelusive and contradictory in many cases. Hovvever. Üıese fındings have been utilized in development and design of educational policies and programs for improving smdent aclıievement in public schools.

Most studies use regression based analysis metiıods in examining the relationship betvveen aclıievement outeomes and various SES and school related variables. As noted by Madaus and others (1980) existence of interrelations among e\planatory variables can result in biased estimation of regression coefficients. Poliey implications dravvn from these studies regarding vvhat constitutes an effective school may be misleading.

Units of Analysis and Data Aggregation

The inconsisteney vvitlıin each study betvveen tlıe aggregation levels of input and output data results in a statistically biased estimation of relationships betvveen dependent and independent variables. Therefore. Glassman and Biniaminov (1981) argue that conelusions derived from studies using various levels of data aggregation should be treated vvitlı caution because tlıe effect of any independent variable may be underestimated. Most often tlıe dependent aclıievement measure is a test score for the individual smdent, but tlıe independent variables are aggregated to tlıe school level. For example, vvhen family background factor variables are aggregated to school level, vvhich is very common in input-output studies, their estimated effects on outeomes relative to school effects are inflated.

This statistical bias is created not only by data aggregation itself. but also by the conceptualization of independent variables. Since most surveys consider only a narrovv' range of school variables, and "schools tend to be more homogeneous than are families. it necessarily means Üıat (as assessed in terms of proportion of population variance accounted for) family variables vvill usuâlly have greater effect than school variables" (Rutter, 1983, p.6).

It is problematic and inappropriate. both staüstically and conceptuaHy. to infer relations about one level from aggregated data at another level. Hovvever. attempts to avoid aggregation bias are most often directed at searclüng for an appropriate level and unit of analysis. For example, Kiesling (1969) argues that using school district aggregates for such variables as persomıel characteristics. conunumty characteristics. ete. "iç similar to using an out-of-focus telescope" (p.3). But he ^gg-^uio üıai <n apprepnate unit of analysis should be Üıe school or wiJ3aroom level. ignoring the hierarchical nature of Üıe educational data. If school level is preferred as tlıe correct unit of analysis. most often. researehers make inferences about individual behavior at Üıe micro-level from school level analysis.

As Burstein (1980) points out, the reason for using aggregate data involves practical concenıs rather than a statistical or conceptual rationale. Because of concenıs of coııfidentiality of individual data. the availability of aggregated data. difficulties and costs associated vvith gathering individual level data and tlıe complexities of analyzing such data. researehers tend to use aggregated data. Hovvever. using aggregated data can be problematic vvhen estimating individual effects and cross-level effects as changes oecur in Üıe meaning of tlıe same variable measured at different levels. For example the same variable measured at the miero (individual) level is likely to have a different meaning at Üıe macro (group) level. and therefore, may lead to different poliey implications as vvell as practical consequences at each level (see Bryk & Raudenbush. 1992: Burstein. 1980: Leeuvv & Kreft. 1986 for a thorough discussion of tlüs issue). Tlüs point vvas illustrated by Lockheed and Longford (1989) in their analysis of data from the Intenıational Evaluation of Aclüevement (IEA). Second Intenıational Mathematics Study in Thailand. VVhen Üıey used an ordinary regression model for aggregated school level data to predict post-test scores from pre-test scores. each point in pre-test scores accounted for .82 points in post-test scores. vvlüle it shovved a value of .69 in multilevel analysis. In their ordinary regression model, the value of pre-test scores is inflated against any other predictor variable ineluded in tlıe analysis. The use of these tvvo different models may result in different conclusions and may have distinet implications for poliey.

Although focusing on school level variables is considered as a strength in school effectiveness research because of its potential for explaining betvveen school variance in süıdent aclüevement (Coluı and Rosmiller. 1987), it can offer liftle insiglıt for understanding the interaction betvveen school characteristics and süıdeııt aclüevement outeomes or vvitlün-school variance. Furthenııore. exclusion of classroom level variables vvhich can have important negative or positive affects on smdent aclüevement. suggests that given profiles and picüıres of effective schools may be missing important components (Bickel. 1990).

Aggregating stıdent data to the school level can also mask differential effects of school related characteristics on different groups of students. In tlüs case, vvithin school variance in süıdent aclüevement is neglected and is contrary to the empirical evidence shovving that student groups vvith different characteristics respond in different vvays to the same set of school characteristics or tlıe same set of iıterventions vvhere ali süıdents in a particular school are assumed to receive the same treatment and slıovv tlıe same reaction to school related variables. Thus, using aggregated data may promote a misleading poliey implication that e\emplaıy schools are eqıally effective vvith different subgroups regardless of their characteristics. (Purkey & Smith. 1983)

The problems created by using nüxed aggregated level data have been avoided in many school effectiveness süıdies in Tlürd World countries. Because "unlike the siüıation in most industrialized countries. aggregate data. such as socioecononüc indices of different administrative districts have not been readily at hand. As a result of having to consüuct original data sets, Üıe individual pupil has had to be used as the first buüding block" (Riddel. 1989. p.486).

Recent developments in school effectiveness research suggest that problems of data aggregation can be avoided by using Hierarchical Linear Models (HLM), or multi level data analysis. In relation to tlüs. Bryk and Raudenbush (1989) argued that data being used in school effectiveness research is hierarcltically structtıred vvhere muit iple levels of data (for example. from the individual student. classroom. school. district level and state level) should be considered in studying school effects because the grovvüı in knovvledge and skills of the individual süıdent. "typically referred to as learniıg, principally takes place in tlıe organizational settings of schools and classrooms. The structural and normative characteristics of such settings and their extemal environments can have substantial influence on the learniıg processes oecurring vviüün them" (Bryk & Raudenbush. 1989. p. 159). Utilization of a multilevel model allovvs for simultaneously analyzing tlıe data at different levels of the school organization vvithout ignoring interaetive relationslüps vvitiıin and across levels rather than using aggregated data to a single assumed appropriate level (Riddel, 1989).

Multilevel Model

Contrary to the underlying assumptions of the earlier iıput output studies. the multilevel model assumes that achievement or other school outeomes cannot be considered as the produets of additive main effects of variables at a given level of the organizational hierarchy. Because tlıe "miero units" of school organization are nested in the larger context, or "macro units", (Bryk & Raudenbush, 1992) and because variables defıned at different levels of the school system can affect school effectiveness through routine iıtra- and inter-level interaetions among combinations of context dependent variables (Aitkin & Zuzovsky, 1992; Raudenbush & Willms, 1991).

Raudenbush and YVillms (1991) argue that effectiveness of educational reform, or of any particular instructional praetice. may depend on the context of the school organization. Therefore. the concept of "interaction effects" assumed by the multilevel model may have important implications for both poliey formulation and implementation. If, "effects of a reform depend on the background of the child or on the organizational context in vvhich it is implemented" (Raudenbush & VVillms. 1990. p.3), then any poliey aimed at improving educational quality or effıciency must take into account both the student characteristics and the organizational context as vvell as the interaetive nature of ali variables vvithin and betvveen levels.

Development of Multilevel Model

Based on the recent vvorks in data analysis procedures that provide a conceptual framevvork and establish statistical techniques necessary for multilevel data analysis (Aitkin & Longford 1986; Burstein. 1980: Burstein & Linn. 1978; Cronbach. 1976; Dempster. Laird & Rubin. 1976: Goldstein, 1986: Lindley & Smith, 1972; Longford. 1987), a number of multilevel statistical cömputer applications have been developed. These statistical softvvare packages inelude: HLM (Bryk. Raudenbush. Seltzer. & Congdon. 1988). GENMOD (Mason. Anderson & Hayat. 1988). VARCL (Longford, 1988). and ML2 (Rabash. Prosser. & Goldstein. 1989). Applications of the multilevel model to educational data represents the greatest challenge to the conelusions of massive efforts of school effectiveness studies undertaken since the early 1960"s.

Despite the advantages of using multilevel data analysis. there are certain issues vvhich need to be adequately addressed regarding its utilization. Available softvvare of the statistical procedures listed above are not user friendly cömputer applications. and their utilization as vvell as the interpretation of obtained results require an advanced level of statistical knovvledge. A technical revievv of four statistical softvvare packages of the multilevel model available (Kreft. Leeuvv, & Kim, 1990) indicates that the Hierarchical Linear Model (HLM) is more advanced than the other three statistical packages in terms of technical capabilities and user friendliness[3].

[3] A discussion by VVilliam M. Mason (1991) in Sociological Methodology, Volume. 21, American Sociological Association. leads to the conelusion that available cömputer programs that can be utilized for the Bayesian statistics are "quite inadequate". Mason calls attention to the need for development of more convenient cömputer programs. so one can use "vvithout investing the rest of (p.347) his/her life in them.

Hierarchical Linear Model

The formulation of HLM is based on Bayesian estimation procedures. The term "hierarchical linear modeL[1] vvas fırst introduced by Lindley and Smith (1972) in their study of Bayesian estimates for hierarchically structured data. vvhich represented the fırst majör breakthrough in establishing a statistical model for multilevel data analysis. Their vvork vvas follovved by Smith (1973) vvho attempted to develop Bayesian linear estimation procedures. Hovvever, these initial studies of Bayesian estimation could not overcome the complexities of Bayesian covariance components estimation. Wide spread utilization of the HLM became.feasible only after further computational developments took place in the late 1970's and the 1980~s. The problems of Bayesian covariance components estimation vvas resolved by the vvork of Dempster, Laird and Rubin (1977) vvhich established a statistical procedure to obtain maximum likelihood estimation for large scale data via EM algorithm[2]. Dempster. Rubin, and Tsutakavva (1981) applied the EM algorithm to a random coefficients regression model. Their application provided the practical computational details for utilization of the EM algorithm in hierarchically structured data settings.

In addition to breakthroughs in estimation of maximum likelihood via EM algorithm, Goldstein (1986) formulated an iterative generalized least squares approach and Longford (1987) developed a Fisher scoring algorithm for maximum likelihood estimation of covariance components in multilevel mixed linear models.

Follovving these developments, Bryk, Raudenbush, Seltzer, and Congdon (1988) formulated a data analytical model, HLM, enabling "researehers to formulate and test explicit statistical models for processes oecurring vvithin and betvveen educational units" (Raudenbush. 1988, p.86). Appropriate error struetures such as random intercepts and random coefficients can be specifıed by using this model, vvhich can not be performed by conventional data analysis procedures

Although most applications of the HLM involves tvvo hierarchical levels, there have been several successml attempts to extend its application to three levels (see Bryk and Raudenbush, 1989 and 1992 for formulation of a three level model). In its simplest form, a tvvo level analysis requires tvvo interrelated equations. The fırst equation represents a vvithin-unit or a miero level analysis formulated as;

$$Yij = J8JO + /3jiX_{y!} + J3_{:z}X_{a}2 + + fijp\text{-}iX_{lip„}i + R,j \quad (1)$$

The term "EM algorithm" vvas fırst introduced by H. O. Hartley, "Maximum Likelihood Estimation From Incomplete Data." Biometrics, 14, (pp. 174-194).

vvhere: $y^\wedge$ represents the outcome for individual / in macro unit/. X_{yp} is the value of the vvithin-unit predictor k for individual / in unit /, and $R,,$ is the random error term. Stnıctural relationships of the Xj_P individual level variables vvith the outcome $y,$ vvithin unit / is captured by coefficients of $/3_{JP}$ vvhich are presumed to vary across macro units. $j3_{!p}$ are regression coefficients that characterize the structural relationships vvithin unit /; for,

/— 1......nj students vvithin school /;

$j=$ 1......A'schools; and

$p=$ 0.....P-l independent variables in the first stage model.

The vvithin unit model assumes that the error Ry are normally distributed vvithin each school vvith a mean of 0 and constant variance ar (residual sampling variance). The model explained in equation (1) is a standard linear model except it allovvs vvithin unit regression coefficients, $P,_P$, to vary aceross macro level units (schools).

In order to take variation across macro units into account. the betvveen unit model formulates the variability in each P structural (regression) parameters, $f3_{JP}$, as a funetion of unit level variables. $@_{Ip}$. and random error, 11_{iv} (Raudenbush and Bryk, 1988, p.434). Then the betvveen-unit model is formulated as:

$$Pip \ - \quad 0()_p + \ OipZı,-+ \ OipZij \ + \qquad + \qquad 0Q.ı_pZg\text{-}lj+ \ \ U/p \qquad (2)$$

structural relations in unit j	***uffect of unit level characteristics on M'ithin-unit relations***	***unigue random effect associated with unit j***

vvehere, $U_{!p}$ represents random error at school level ;

Zqj are values of the school level (betvveen unit) variables for school /; for.

q = 0......Q-l independent variables in the second level model; and,

6_{qp} are the regression coefficients that capture the effects of school-level variables on the vvithin-school structural relationships, $J3_{!P}$.

Bryk, Raudenbush. Seltzer and Congdon (1988) notes that this tvvo level model enables us to achieve several objeetives:

1. Because the model permits estimation of both an average vvithin-school and a betvveen-school regression equation, vve can decompose any observed relationship into its betvveen- and vvithin school components.
2. VVe have a multivariate formulation for examining the effects of betvveen-group factors (e.g.. school policies and practices) on vvithin-school phenomena (e.g., the average achievement and SES-achievement relationship.

3. We can estimate vvithin school regression coefficients, $J3_{ip}$, that are udjested for other confounding variables vvithin-school.
4. The estimated slopes. $J3_{ip}$, are vveighted in proportion to their precision in the regression against school level factors. Precision is also enhanced by the fact the estimation of 0_{qp} utilizes information on the correlation among the estimated multiple vvithin-school regression coefficients, $J3_{jp}$.
5. We are able to generate better estimates for the vvithin-school structural parameters, $J3j_{p}$. than are available through a traditional regression model vvhich only uses the data from school *j.* As aresult vve can arrive at a better descriptive characterization of each school that mighl: be usefül, for example, in research that seeks to identify usually effective schools (P- 5).

Misestimation of effects is alvvays a potential problem in conventional educational data analysis procedures. Because individuals are not randomly assigned to groups and individuals are nested in classrooms, classrooms in schools, and schools are nested in a larger district organization. An explicit assumption in equation 2 presented above is that the effect of unit level characteristics on vvithin-unit relationships varies as a function of contextual factors associated vvith each unit. Therefore, there is a unique random effect associated vvith each unit (Braun 1989; Bryk and Raudenbush, 1992).

Use of various estimations in decision making process is not an uncommon issue in praetice. Graduate schools very often justify their admission decisions by predietion of academic success from standardized test scores of their applicants. Hovvever, Braun, Jones, Rubin, and Thayer empirically challenged the validity of the predietion itself used in decision making throughout tlıe admission process. They studied the predietion of academic success (FYA) from GMAT test scores in 59 business schools. Tvvo different GMAT scores, GMAT-V, and GMAT-Ç), vvere utilized as predictors vvhich posed some problems. Hovvever, the most serious problem arose from unbalanced distribution of students and applicants, and varying effects of contextual variables in predietion. Predietion of FYA from GMAT for minority students vvould be clearly biased in conventional data analysis procedures. since available data is primarily dominated by vvhite students. Because there vvas no minority enrollment in 14 schools, and there vvere only one to three black students in 20 business schools out of a total sample of 59 schools. Since, predieting minority academic aclıievement from overall data set vvas not a credible procedure, another alternative vvitlıin tlıe framevvork of conventional models vvould be a separate equation by race. Hovvever. given tlıe sparse nature of data, modeling a separate equation for minorities vvould not be a feasible approach to solve this problem.

Braun et al. (1983) utilized a multilevel model approach and formulated separate equations for each school for minorities and vvhites by using information from a vveighted composite estimator from each school, and the relations that exist in the total sample (see Dempster, Laird. and Rubin,

1976 for details of computational procedure used by Braun et al. for estimation). Their analysis indicated that prediet ion coefficients for minority students vvere significantly different from coefficients for vvhites. Further, prediet ion of academic achievement varied depending on organizational characteristics of each school. Therefore, existing alternatives vvithin conventional data analytical procedures vvould have predieted a misestimated. or biased FYA score for minority applicants

As illustrated in the business schools example, HLM enables the researeher to "formulate and test hypotheses about hovv variables measured at one level affect relations oecuring at another" (Bryk and Raudenbush, 1992, p.6), e.g., hovv organizational characteristics affect relations betvveen GMAT scores and academic achievement. In most cases, researeher s very often are interested in effects of policies or practices at classroom, school, district. and system level on student achievement or behavior. Examination of such cross-level effects by traditional methods of data analysis raise serious doubts about validity of cross-level inferences and as a necessity imposed by the nature of data itself, an accurate assessment of effects requires a multilevel model.

Tlıe study of early vocabulary grovvth in children by Huttenlocher, Haiglıt. Bryk, and Seltzer (1991) illustrated this issue by utilizing HLM to assess effects of gender and exposure to language during infaney. They established an individual vocabulary grovvth trajectory for each child at Level I. Grovvth parameters at Level I vvere predieted by a set of variables at Level II such as amount of maternal speech and child's sex. The effects of maternal speech on a child's vocabulary grovvth in this study vvas much greater than tlıe conventional model estimates. Because HLM analysis provided a more accurate estimate of effects by using data from each repeated observation rather than using mean score of observations or pre test-post test scores for estimation.

Dravving on the development of EM algorithmic approach (Dempster. Laird, and Rubin, 1977), the basic model of HLM presented by equation I and equation II partions variance into vvitlıin- and betvveen group components. Aitkin and Longford (1986) demonstrated that effectiveness ranking of educational institutions in conventional studies of school effectiveness can be misleading. VVhen student level variables aggregated to school level. "pupil level variables can reduce the school level variance component, if the mean of the variables varies oyer schools" (Aitkin and Longford, 1986. p.15). Bernstein (1990) provided very conelusive evidence supporting their point in an attempt to test and establislı a predietive model' of student aclıievement in Pennsylvania school districts. VVhen he tested the effectiveness of the HLM model against the OLS model, his analysis indicated that the HLM model accounted for 62% of tlıe betvveen district variance in comparison to 51% explained variance by the OLS model.

Conclusions

Problems arise in analysis of data collected at multiple levels. Educational organizations are hierachically structured. Classrooms are nested in schools and schools are nested in larger districts. Conventional data analysis techniques, for example, OLS analysis, assume variability of each variable is identical. Hovvever, results of multilevel analyses indicate that this is not the case. Variability of one variable at classroom level is very different from its variability at school level. In school effectiveness studies, this may be true of family background variables, resources inputs and instructional materials.

Hierarchy is a fundamental characteristic of educational organizations and educational data. Aggregation bias occur if data are aggregated to the group level, ignoring vvithin group variation, or analyzed solely at the student level, ignoring group effects. The multilevel analysis allovvs for explicit modelling of effects at both levels so that ali estimated effects are adjusted both for individual level and group level influences on the outcome.

References

Aitkin, A. & Longford, N. (1986). Stattistical modeling issues in school effecttiveness studies. *Journal öf Royal Statistical Society,* **Series A, 149, 1-43.**

Aitkin, M. and Zuzovsky R. (1991). Curricular change and science achievement in Israeli elementary schools. in Stepnen W. Raudenbush and J.D. VVillms (eds.) *Schools, classrooms, and pupils: International studies of schooling from a multilevel perspeetive,* **Nevv York: Academic Press.**

Aitkin, M. and Zuzovsky R. (1992). *New paradigm for the analysis of hierarchically structured data in school effectiveness studies.* **Paper presented at the American Educational Research AssociationAnnual Meeting, San Francisco.**

Bernstein, L.S. (1990). *The application of hierarchical linear modeling to multilevel student achievement data,* **Ph.D. Dissertation, University of Pittsburgh:**

Bickel, W.E. (1990). The effective schools literatüre: **implications for research and praetice. In** *Handhook of School Psychology.* **Nevv York, VViley & Sons.**

Braun, H.I.. Jones. D.H.. Rubin, D.B., & Thayer, D.T. (1983). Empirical bayes estimation of coefficients in the general linear model from data of deficient rank. *Psychomethka,* 489, (2), 171-181.

Bryk, A. S. Raudenbush, S.W., Seltzer, M., & Congdon, RT. (1988). *An introduction to HLM: Computer program and ıısers guide.* University of Chicago, Department of Education.

Bryk. A.S. and Raudenbush, S.W. (1992). *Hierarchical linear models: Applications and data analysis methods.* Nevvbury Park: Sage. -

Bryk, A.S.. & Raudenbush, S.W. (1989). Tovvard a more appropriate conceptualization of research on school effects: A tlıree-level hierarchical linear model. In R.D. Bock (Ed.), *Multilevel analysis of educational data* (pp. 297-310). Orlando, FL: Academic Press.

Burstein, L. (1980). The analysis of multilevel data in educational research and evaluation. *Revievv of Research in Education,* 8,158-233.

Burstein, L.. Linn. R.L.. & Capell, F.J. (1978). Analyzing multilevel data in the presence of heterogenous vvithin-class regressions. *Journal of Educational Statistics,* 3. 347-383.

Burstein, L.; Kim, K.S. and Delandshere G. (1989). Multilevel investigations of systematically varying slopes: issues, alternatives, and consequences. in R. Darrel Bock (ed.) *Multilevel analysis of educational data,* Nevv York: Academic Press.

Cohn, E., & Rossmiller, R. (1987). Research on effective schools: implications for less developed countries. *Comparaüve Education Revievv,* 31(2), pp. 377-399.

Coleman, J., Campbell, E.. Hobson. C, McPartland. J.. Mood. A., VVeinfall. F., & York, R. (1966). *Equality of educational opportunity.* Washington, DC: Department of Health, Education and vVelfare.

Cronbach, L.J. (1976). *Research on classroom and schools: Formulations of questions design and analysis.* Occasional paper. Stanford, CA: Stanford Evaluation Consortium.

de Leeuvv, J. & Kreft, I. (1986). Random Coefficient models for multilevel analysis. *Journal of Educational Statistics,* 11, (1), 57-85.

Dempster, A. P.. Laird, N. M., & Rubin. D. B. (1977). Maximum Likelihood from incomplete data via the EM algorithm. *Journal of Royal Statistical Society,* Series B, 39, 1-8.

Glasman, N.S. and Biniaminov, 1.(1981). Input- output analyses of schools. *Revievv of Educational Research,* 51(4), 509-539.

Goldstein, H. (1989). Models for multilevel response variables vvith an application to grovvth cuıves. in R. Darrel Bock (ed.) *Multilevel analysis of educational data,* Nevv York: Academic Press.

Goldstein, H.I. (1986). Multilevel mixed linear model analysis using iterative generalized least squares. *Biometrika,* 73 (1). 43-56.

Huttenlocher, J. A., Haight, W.. Bryk. A.S.. & Seltzer, M. (1991). Early vocabulary grovvth: Relation to language input and gender. *Developmental Psychology,* 27 (2), 236-249.

Keeves. J.P. & Sellin, N. (1990). Multilevel analysis. In H.J. Walberg & G.D. Haertel (Eds.), *The International Encylopedia of Educational Evaluation.* Pergamon Press.

Kiesling, H.J. (1969). *The relationship of school inputs to pııhlic school performance in New York State.* ED #051 585.

Kreft, I.G., de Leeuvv, J., & Kim, K. (1990). *Comparing four statistical packages for hierarchical linear regression: Genmod, HLM, ML2, and VARCL.* Los Angeles: UÇLA.

Lindley, D.W., & Smith, A.F.M. (1972). Bayes estimates for Linear model. *Journal of Royal Statistical Society,* Series B, 34, 1-41.

Lockheed, M.E. and Longford, N.T. (1989). *A multilevel model of school effectiveness in a developing country,* VVorld Bank Discussion Papers, The VVorld Bank, Washington, D.C.

Lockheed, M.E. and Longford, N.T. (1991) School effects on mathematics achievement gain in Thailand. in Stephen W. Raudenbush and J.D. Willms (eds.) *Schools, classrooms, and pupils: International studies of schooling from a multilevel perspeetive,* Nevv York: Academic Press.

Longford, N.T. (1987). A fast scoring algorithm for maximum likelihood estimation in unbalanced mixed models vvith nested random effects. *Biometrika,* 74 (4), 817-827.

Longford, N.T. (1988). Fisher scoring algorithm for variance components analysis of data vvith multilevel strueture. In R.D. Bock (Ed.), *Multilevel analysis of educational data* (pp. 297-310). Orlando, FL: Academic Press.

Madaus, G.F., Airasian, P.W., and Kellaghan, T.(1980). *School effectiveness: A reassessmentofthe evidence.* N.Y.: McGraw-Hill.

Mason, W.M., Anderson, A.F., & Hayat, N. (1988). *Manual for Genmod.* Ann Arbor: University of Michigan, Population Studies Center.

Mislevy, R.J. and Bock, R.D. (1989). A hierarchical item-response model for educational testing. in R. Darrel Bock (ed.) *Multilevel analysis of educational data,* **Nevv York: Academic Press.**

Moorhead, R. and Nediger, W.G. (1990). *Process components of effectiveness: A holistic study of 4 secondary school principals.* **Paper presented at the Annual Meeting of the Canadian Society for the Study of Education, Victoria, British Columbia. (ED. 323 661)**

Purkey, S., & Smith, M. (1983). Effective schools: A reviewv. *Elementary School Journal,* **83(4), 427-52.**

Rabash, J., Prosser. R., & Goldstein, H. (1989). *ML2: Softvvare for hvo-level analysis. Users' guide.* **London: University of London, Institute of Education.**

Raudenbush, S. W. (1988). Educational applications of hierarchical linç models: A reviewv. *Journal of Educational Statistics,* **13 (2), 85-116**

Raudenbush, S.W. and Bryk, A.S. (1989). ^ovvard a mcre appropriate conceptualization of research on school effects: A three-level hierarchical linear model, in R. Darrel Bock (ed.) *Multilevel analysis of educational data,* **Nevv York: Academic Press.**

Raudenbush. S.W. and VVillms, J.D. (1991). The organization of schooling and its methödological implications. in Stepnen W. Raudenbush and J.D. VVillms (eds.) *Schools, classrooms, and pupils: International studies of schooling from a multilevel perspeetive,* **Nevv York: Academic Press.**

Riddell, A. (1989). An alternative approach to the study of school effectiveness in Third VVorld countries. *Comparative Education Review.* **33, 4 (Nov) 481-506.**

Rutter. M.(1983). School effects on pupil progress: Research fındings and poliey implications. In Shulman. L.S. and Sykes. G.(eds), *Handbook of teaching and poliey.* **Nevv York: Longman.**

Seltzer, Micheal H. (1995). Furthering our understanding of the effects of educational programs via a slopes as-outeomes framevvork. *Educational Evaluation and Poliey Analysis.* **17, 3: pp. 295-304.**

Author

Dr. Emin Karip is assistant professor at Gazi University, Gazi Education Faculty, Department of Educational Sciences, Beşevler, Ankara. His areas of specialization are educational administration. quantitative analysis in educational poliey and school improvement.