DOCUMENT RESUME

ED 356 946 | SE 053 044

AUTHOR	Young, Deidra J.; Fraser, Barry J.
TITLE	School Effectiveness and Science Achievement: Are There Any Sex Differences?
PUB DATE	Mar 92
NOTE	24p.; Paper presented at the Annual Meeting of the National Association for Research in Science Teaching (Boston, MA, March 21-25, 1992). For a related article, see SE 053 045.
PUB TYPE	Reports - Research/Technical (143) -- Speeches/Conference Papers (150)
EDRS PRICE	MF01/PC01 Plus Postage.
DESCRIPTORS	*Academic Achievement; Correlation; Cultural Background; Databases; Family Influence; Foreign Countries; Mathematics Achievement; *Performance Factors; *School Effectiveness; Science Education; Secondary Education; Secondary School Students; *Sex Differences; Socioeconomic Influences; Student Attitudes; *Student Characteristics; Verbal Ability
IDENTIFIERS	Australia; *Science Achievement; *Second International Science Study

ABSTRACT

The purpose of this study was to use secondary analysis of an Australian database known as the Second International Science Study (SISS) to examine the role of student, school and home factors in explaining student differences in science achievement, especially on sex differences. Student characteristics investigated in the study included home background, attitude towards science, ethnic background, verbal and mathematics abilities, and sex of the student. The study applied the Hierarchical Linear Model (HLM) to analyze data from 4,917 14-year-old students from 233 Australian schools; it also incorporated the Rasch analysis model for the construction of the science achievement scale. The correlations between student characteristics and their science achievement indicated that sex, attitude towards science, and ethnicity of the student are comparatively weak correlates of science achievement and that verbal and quantitative ability and socio-educational level had much stronger correlations. School effects found to be statistically significant in influencing student differences in science achievement included the percentage of female teachers in the school, teacher decision making, average socio-educational level of students, average verbal ability, average quantitative ability, and average ethnicity of students. These student and school effects were utilized to develop a preliminary integrated model of school effectiveness. No statistically significant variations in sex differences in science achievement between schools was found. The study concluded that results of this preliminary study indicate that further path analysis and use of the HLM for all countries in the SISS should provide useful information regarding the relationship between student achievement, attitudes towards science, and other home and school environment characteristics. A list of 29 references is provided. (MDH)

School Effectiveness and Science Achievement: Are there any Sex Differences?

Deidra J Young and Barry J Fraser
Science and Mathematics Education Centre
Curtin University of Technology, Perth, Western Australia

Paper presented at the Annual Meeting of the
National Association for Research in Science Teaching
21-25 March 1992, Cambridge Hyatt Regency Hotel, Boston, Massachusetts

This research was supported by a Research Fellowship from the Australian Research Council, the Australian national Key Centre for School Science and Mathematics (Especially for Women) and the Science and Mathematics Education Centre, (both located at Curtin University of Technology, Perth, Western Australia). The authors are grateful for the consultative assistance provided by Dr Raymond Adams, Chief Research Officer, Australian Council for Educational Research, Melbourne, Victoria, Australia. In addition, we would like to thank Dr Adams for generously providing the Rasch analysis package, TITAN, and for his assistance in the use of this package. Finally, we would like to acknowledge the technical advice of Professor David Andrich who so ably taught the first author about Georg Rasch and the Rasch Model.

School Effectiveness and Science Achievement: Are there any Sex Differences?

Deidra J Young and Barry J Fraser
Science and Mathematics Education Centre
Curtin University of Technology, Perth, Western Australia

Introduction

The purpose of this study was to use secondary analysis of an Australian database known as the Second International Science Study to examine the role of student, school and home factors in explaining student differences in science achievement, especially sex differences in achievement. This study employed methodology which accommodated both the Rasch model (science test items were calibrated for item difficulty and student ability) and the multilevel nature of the data (students clustered within schools).

Student characteristics investigated in this study for their influences on science achievement included the socio-educational level (home background), attitude towards science, ethnic background, verbal and mathematical abilities and sex of the student. School effects were also investigated in terms of their ability to explain student differences in science achievement. These included aggregated student characteristics, the school type (private /public), the sex composition of the school (single-sex/coeducational schools), the local population, the ratio of students to teachers, the percentage of female secondary teachers, the degree of teacher decision making, the involvement of the parent-teacher committee and the source of operating costs for the school. The school effects were categorized as contextual or process depending on their nature. Contextual school effects reflected the school's physical resources, student-body composition and other school environmental factors, whereas the process school effects are related to the management and organizational structures of the school. In this paper, school effectiveness research is reviewed and then related to an integrated model of school effectiveness. Following this, the sample is described and methodology summarized. Finally, the relevance and importance of this study are elaborated.

School Effectiveness Research

The central proposition of this paper is that school characteristics and the school environment can influence student achievement and sex differences in science achievement, and this cannot be ignored by researchers in the field of gender research. Previously, there has been considerable research emphasis on the ability and family background of the student in determining academic performance. The Coleman Report (Coleman et al., 1966, p. 296) estimated that the percentage of school influence on student achievement was about 10 to 20 percent of the total variance, yet the methodology used by Coleman did not account for the hierarchical nature of students nested in schools. Further, Coleman's study found that the American public schools were remarkably homogeneous both in terms of educational resources and student outcomes. Unfortunately, this study also demonstrated that large differences in student achievement existed between major racial and socioeconomic groups. Coleman's findings were repeated in further large-scale studies (Jencks et al., 1972, 1979; Hauser, Sewell & Alwin, 1976), with the resulting school effectiveness research finding that (1) school level variables account for relatively small amounts of variance in student achievement, (2) physical resources at the school level account for relatively small amounts of the variance between schools and (3) student characteristics such as socioeconomic status and home background should be used to adjust student achievement in statistical analysis of such large-scale studies (Scheerens, 1990).

While Coleman's research focused on the small effect which school characteristics have on student achievement when home background is controlled, more recent school effectiveness research has attempted to open the black box of the school environment and examine more closely what happens in schools. The current school effectiveness research genre has shifted focus from the black box to opening up within-school characteristics which include school organization, culture and educational technology. Case study research has become more useful in this respect, as well as large-scale studies such as the International Association for the Evaluation of Educational Achievement (IEA) science studies: First International Science Study (Comber & Keeves, 1973) and Second International Science Study (Postlethwaite & Riley, 1992). This paper describes some analyses of data from the Australian database of the Second International Science Study (SISS) collected in 1984. In the recent school effectiveness research, typified by Mortimore et al. (1988), a combined large-scale study and case studies, which found that student progress was enhanced by effective schools irrespective of the student's own social class, sex or ethnic group. Moreover, Mortimore asserted that:

> Even though overall differences in patterns of pupil attainment are not removed in the most effective schools, the performance of all children is raised and, as we have demonstrated, disadvantaged children in the most effective schools can end up with higher achievements than their advantaged peers in the less effective schools.
> (Mortimore, 1988, p. 217)

Scheerens and Creemers (1989) have criticised the following five factor model of effective schools' characteristics derived from recent research:

1. Strong educational leadership,
2. High expectations of student achievement,
3. Emphasis on basic skills,
4. A safe and orderly climate, and
5. Frequent evaluation of pupils' progress.

These researchers assert that this model is correlational only, not causal, in that the effects of high achievement could lead to a safe and orderly climate and not vice versa. Emphasis on basic skills is also a questionable goal because higher-level cognitive learning and affective outcomes are equally important. Finally, the independence of these five factors is doubtful; frequent evaluation and orderly climate are likely to be a reflection of strong instructional leadership (Scheerens & Creemers, 1989, pp. 692-693).

Scheerens (1990, p. 63) proposed an analytical scheme (Figure 1) to explain indicator systems, which incorporates the contextual data (policy, school environment), inputs (resources, teacher characteristics), process data (curriculum, school organization and climate), ouputs (achievement) and outcomes (employment and income). The use of process indicators as an explanation of why some schools do better than others, in terms of student achievement, should not be ignored in school effectiveness research contends Scheerens.

An integrated model of school effectiveness developed by Scheerens (1990, pp. 72-73) was used in order to define school- and student-level educational performance indicators. This model incorporates an analytic systems model which recognises context, input, process and output variables; a multi-level framework discriminating between student and school characteristics; perspectives for conceptualising the interrelationships between variables defined at different levels (most notably contingency theory and organisational theory that incorporates the schools' primary processes); and substantive findings from different types of educational effectiveness research (see Figure 2).

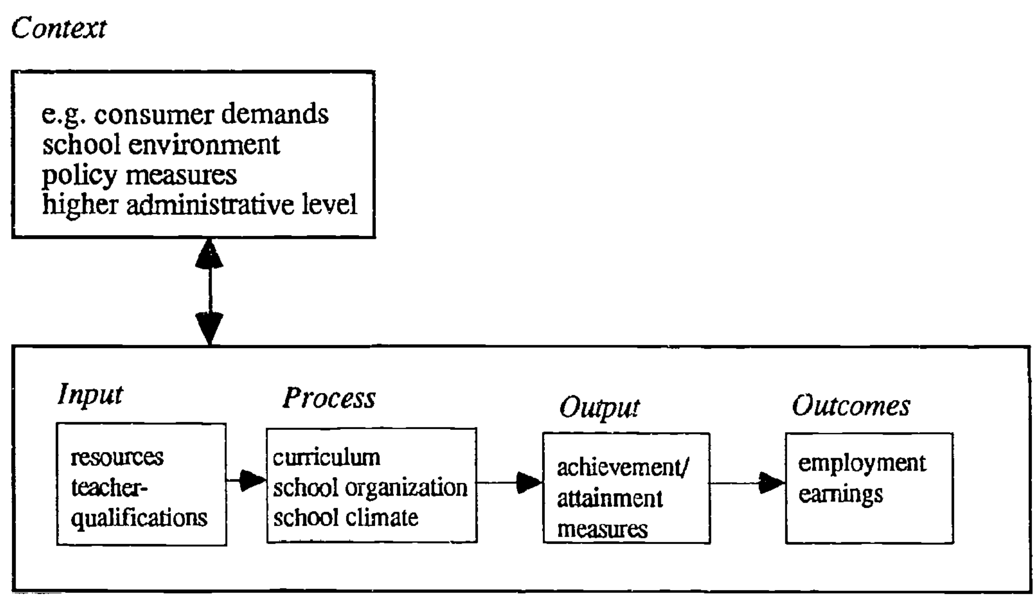

Figure 1. Context-input-process-output-outcome Model of Schooling
(Scheerens, 1990, p. 63)

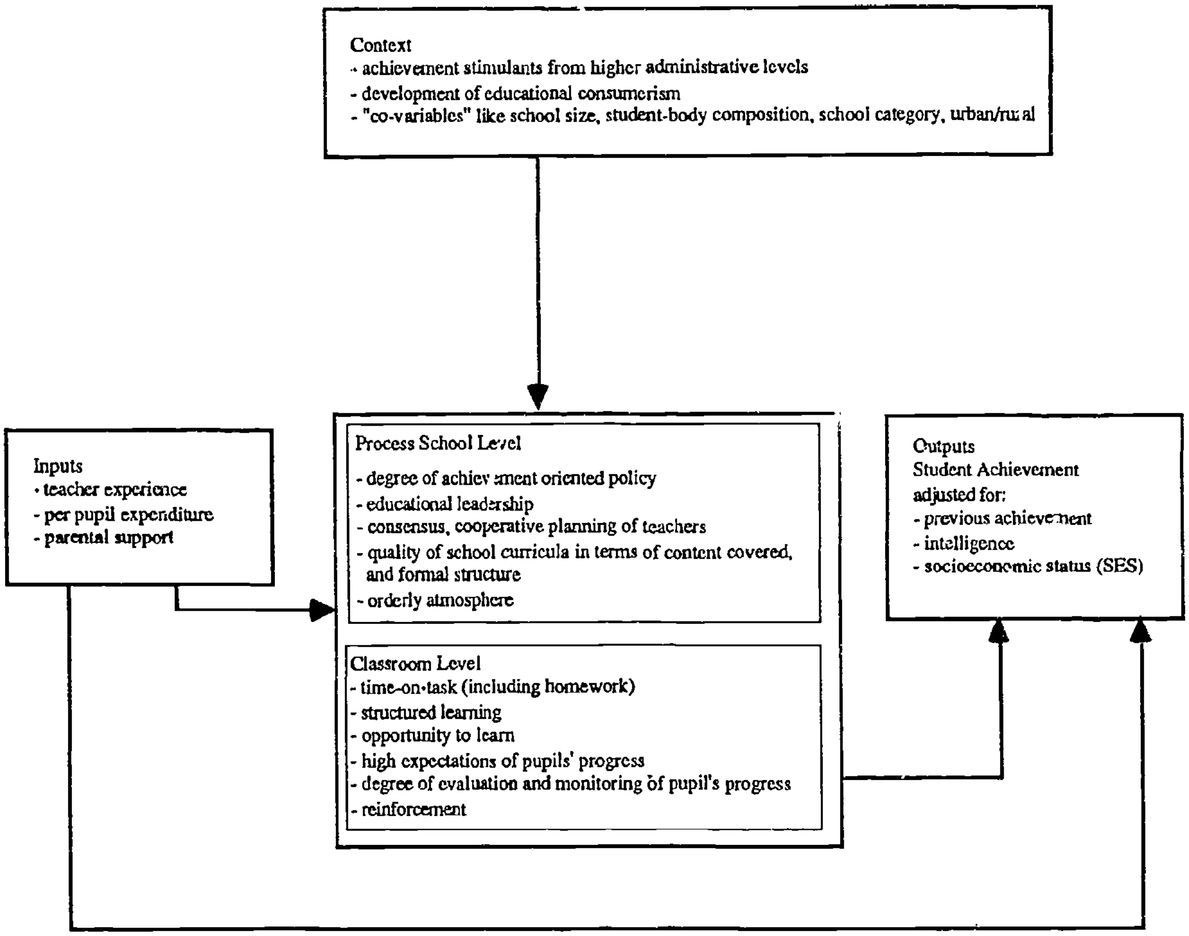

Figure 2. Integrated model of school effectiveness.
(Scheerens, 1990, p. 73)

Objectives and Theoretical Framework

The purpose of this study was to investigate the role of student, school and home background factors in explaining differences in science achievement using a database from the Second International Science Study. Additionally, sex differences in science achievement were examined with respect to other student and school characteristics.

Although sex differences in science achievement have been investigated in previous large-scale studies (Keeves, 1973; Comber & Keeves, 1973; Kelly, 1978), some of these studies assumed that the data had been collected as a simple random sample and therefore ignored the stratification of the sample by state/territory and school type (government, Catholic and independent), while others neglected the multilevel nature of data from students who are clustered within school systems. This study employed methods of analysis which accommodated both the complex sample design and the multilevel nature of the data.

The integrated model of school effectiveness (Figure 2) suggested by Scheerens (1990) includes a contextual effect, teacher and parental inputs, school and classroom process effects, and student achievement outputs. The student achievement outputs are adjusted for previous achievement, intelligence and socioeconomic status. In this study, some of the contextual effects examined included the size of the population in the school area, the ratio of students to teachers and the percentage of female secondary teachers in the school. Schools were categorised into three school types: government, Catholic or independent. However, in this study, the variable school type was coded in the same way for both Catholic and independent schools and differently for government schools. The student body composition was characterised by sex composition or single-sex versus coeducational, average socio-educational level, average verbal ability, average quantitative ability, average ethnicity and average attitude towards science. Included in the school context was the source of operating costs as either private or public. Two process-school level components in this study were the parent and teacher committee's involvement in the running of the school and a measure of teacher decision making in the school. These variables are defined more fully in Table 1.

The output measure of student achievement for this study was a science achievement, a scale constructed from 46 science test items. All of the test items were multiple-choice in format and covered the three science content areas of biology, chemistry and physics. In addition, the items varied in the degree of cognitive ability required (from concrete to analytical). The science achievement output of students was adjusted for the following student characteristics: sex, attitude towards science, ethnicity, verbal ability, quantitative ability and socio-educational level. In summary, the integrated model of school effectiveness investigated in this study included the following components:

Type of Indicator	*Variable Description*	*Variable Name*
Context	Size of population of school area	areapopn
	Student/teacher ratio	stutchr
	Percentage of female teachers	femtch
	Average socio-educational level	avsel
	Average verbal ability	averb
	Average quantitative ability	avmath
	Average ethnicity	minority
	Average attitude towards science	avenjsci
	School type	schtyp
	Sex composition of the school	sexcomp
	Source of operating costs	opcost
Process	Teacher decision making	tchdec
	Parent teacher committee involvement	ptc

Output	Science achievement	science
	Sex	sex
	Attitude towards science	enjsci
	Ethnicity	ethnty
	Verbal ability	tot2v
	Quantitative ability	tot2q
	Socio-educational level	sel

While these school contextual and process effects on student performance in science are somewhat limited when compared with the Scheerens model (Figure 2), this paper represents only a preliminary report. Further research into inputs, context, process and outputs using the Second International Science Study database will include teacher characteristics and school expenditure.

Description of Sample

There were 4917 14-year-old students in this sample (2565 girls and 2352 boys) selected from 233 schools. The target population of 14-year-old Australian students consisted of 246 132 students within 2144 schools at the time of this survey (1983/84). The sample design used in this study was a stratified two-stage cluster design, with schools selected randomly from within each of 24 strata (the eight Australian states and the three school types: government, Catholic and independent) and students selected randomly from within each school. Table 2A presents a listing of the 24 strata and a description of the schools in each strata. The actual sampling statistics are found in Table 2B, with each strata presented by targeted population sizes, designed sample sizes and achieved sample sizes. The sampling fractions, percentage response rate and stratum weights are also provided here. This complex sample design meant that the normal assumptions of simple random sampling could not be made validly in order to test statistical significance. For this reason, a multilevel model was developed which accounted for the nested nature of the data. In addition, students in each strata were weighted according to the proportion of students whom they were supposed to represent.

Methodology

The Hierarchical Linear Model

The use of powerful computers to analyse large databases during the First International Mathematics Study focused on the home and school variables influencing student achievement, while using the student as the unit of analysis (Husén, 1967). However, this type of analysis did not adequately address the variability of schools contributing to statistical tests of significance. Studies in the United States tried to compensate for this problem by only looking at the school differences; for example, Coleman et al. (1966) used the school as the unit of analysis. Further, a large database was analysed by Peaker (1967) in England by examination of between-school means (aggregated student data) and pooling between students. However, these studies ignored the differences between-students within- schools which may contribute towards explaining the variance. The controversy over the most appropriate unit of analysis has continued until the importance of a different approach to educational research was first proposed by Cronbach:

> *The majority of studies of educational effects - whether classroom experiments, or evaluations of programs, or surveys - have collected and analysed data in ways that conceal more than they reveal. The established methods have generated false conclusions in many studies.* (Cronbach, 1976, p. 1)

Most educational research revolves around students who receive schooling in classrooms located within schools, within school districts, within states, etc. The grouping of students, classes and schools occurs in a hierarchical order with each group influencing the members of the group in thought and behaviour. The nature of these hierarchical structures produces multilevel data. Theories about the effects of the multilevel structure of education (the different levels of the educational hierarchy) should lead to attempts to specify models which involve the analysis of multilevel educational data. Burstein (1980) believes that these theories eventually will replace experimental design and analysis with the natural design and analyses that evolve from the multilevel structure of data.

The amount of variation in estimates of variables affecting academic achievement across different levels of analysis cannot be ignored by serious educational researchers. Traditional linear models on which most researchers rely require the assumption that subjects respond independently, yet most subjects are 'nested' within classrooms, schools, districts, states and countries so that responses within groups are group dependent. To ignore the nested structure of this type of data ultimately will give rise to problems of aggregation bias (within-group homogeneity) and imprecision (Raudenbush, 1988).

This Hierarchical Linear Model (HLM or multilevel analysis) provided an integrated strategy for handling problems such as aggregation bias in standard error estimates and erroneous probability values in hypothesis testing of school effects. For this study, HLM was chosen as the model most appropriate to study school and student effects relating to science achievement and HLM2 (Bryk, Raudenbush, Seltzer, & Congdon, 1989) was selected as the computer package most suited to analyse the large amount of data in SISS. The use of the HLM in order to investigate the influence of the organisational structure of the school on student performance has been documented by Bryk and Raudenbush (1989, pp. 159-204), Lee and Bryk (1989) and Raudenbush and Bryk (1986). The present study sought to examine the role of school effects in explaining science achievement and sex differences in science achievement. Research on school effects is described as a set of data analysed at the individual student level, with the assumption that classrooms and schools affect students equally. However, when the effects vary among individuals and their contexts, this type of statistical analysis can be misleading (Bryk & Raudenbush, 1987). Ordinary least squares analysis provides information about the total variance, but does not break this total variance into the between- and within-classroom effects. This study endeavoured to explain variations in student outcomes by first decomposing observed relationships into between- and within-school components.

The analysis of gender differences in science achievement and attitudes between different school types must take into account variations from school to school unless inferences are to remain doubtful. Raudenbush and Bryk (1986) point out the fallacies of research findings which ignore the potential effects of the school or classroom as sociological units, citing many research studies with doubtful inferences. These researchers introduced the concept of the hierarchical linear model (HLM), which accommodated both school and student level differences. Their reanalysis of data from a random sample of United States high schools illustrates technical and conceptual advances facilitated by HLM and showed that the relationship between socioeconomic status and mathematics achievement varied substantially across US high schools and that much of this variation was attributable to school type (public versus Catholic). Distinguishing between micro parameter variance (such as school or classroom) and the sampling variance was possible with HLM, which makes it possible to partition the socio-educational level effect into within- and between-group components which yielded an estimate of the school type effect substantially different from earlier estimates. Similarly, Lee's (1986) reanalysis of data from High School and Beyond study (Coleman et al., 1966; Haertel et al., 1987) revealed that differences between public and Catholic schools were attributable to the curriculum and the discipline policies of the schools.

Rasch Analysis

In addition to using multilevel analytic techniques, this study also incorporated the Rasch analysis model for the construction of the science achievement scale (Rasch, 1960). Rasch's models for measurement are based upon Thurstone's models (1925, 1926, 1927, 1959) which were independent of the data. The Rasch model, was further developed into a function of the difference between parameters β_v for person ability and δ_ℓ for item difficulty ($\beta_v - \delta_\ell$):

$$\Pr\{\text{positive response}\} = \frac{\exp(\beta_v - \delta_\ell)}{[1 + \exp(\beta_v - \delta_\ell)]} \qquad \text{Equation 1}$$

For each student, 46 science achievement items were selected and a Rasch calibration produced an ability estimate for all 46 items (based upon equation 1). The 46 science test items consisted of 20 common core test items answered by all students and 26 rotated test items responded to by different subgroups of students. The purpose of the Rasch analysis was to make the science achievement scores independent of the sample and the item difficulty; some of the rotated test items may have been more difficult than others. The science test items also represented biology (20 items), chemistry (12 items) and physics (14 items) content areas. The calibration was made using the Rasch model, which is independent of student ability and item difficulty, and TITAN (Adams & Toon, 1991), a Rasch analysis computer software package for calibrating items according to equation 1.

In Table 3, the frequency of correct scores, item *infit meansquares* and *t* statistics are provided. In particular, large *infit meansquares* (greater than 1.00) are indicative of a misfitting item. This table shows that there were no items which misfit i.e. there were no items much greater than 1.00. Further statistical results are provided in Tables 4 and 5. Table 4 provides a summary of item estimates and case estimates. Note that the reliability of the case estimate was 0.82. The graph of item estimates on all students is presented in Table 5 depicting the spread of items and students. This graph reveals that there is probably a lack of more difficult items (towards the upper part of the graph) in this set of science test items.

Results and Discussion

The correlations between student characteristics and their science achievement is described in Table 6 and reveals that sex, attitude towards science and ethnicity of the student are comparatively weak correlates of science achievement, although statistically significant. The other three student level variables, verbal and quantitative ability and socio-educational level, had a much stronger relationship which was also significant. When the variance in student science achievement was decomposed into school (between-school) and student (within-school) components using the hierarchical linear model (HLM), 15.2 percent of the total variance was found to be related to between-school differences, while 84.8 percent was related to within-school differences (Table 7). The total between-schools variance of 0.179 and within-schools variance of 0.998 resulted in an intra-class correlation of 0.152.

The partitioning of variance in science achievement among students into the within- and between-school components was achieved using the HLM2 computer package (Bryk, Raudenbush, Seltzer & Congdon, 1989). A random average science achievement estimate was specified for the within-school model:

$$Science_{ij} = \beta_{0j} + R_{ij}$$

where $i = 1, \ldots, n_j$ students in school j, $j = 1, \ldots, J$ schools, $Science_{ij}$ represents science achievement of student i in school j, β_{0j} represents the mean science achievement for school j and R_{ij} represents random error of student i in school j.

An unconditional between-school model was also specified:

$$\beta_{0j} = \mu_j + U_j$$

where β_{0j} represents the mean science achievement for school j, μ_j is equal to β_{0j} and U_{ij} represents random error of school j.

The hierarchical linear model was used to include student-level variables to the above model and examine the reduction in unexplained variance. The addition of *sex* to the model reduced the within-school variance by 1.7 percent, while the addition of all of the student characteristics described previously reduced this variance by 43.4 percent (a further 41.7 percent). While sex, ethnic and social differences were notable in their effects on student science achievement, adjustments for these effects were made and then school effects added to the hierarchical linear model.

$$Science_{ij} = \beta_{0j} + \beta_{1j}Sex_{ij} + \beta_{2j}Enjsci_{ij} + \beta_{3j}Ethnty_{ij} + \beta_{4j}Tot2v_{ij} + \beta_{5j}Tot2q_{ij} + \beta_{6j}Sel_{ij} + R_{ij}$$

where *Science* represents student science achievement, *Sex* represents sex of the student, *Enjsci* represents the student's attitude towards the science, *Ethnty* represents the student's ethnic background, *Tot2v* represents the verbal ability of the student and *Totq* represents the quantitative ability of the student and *Sel* represents socio-educational level of the student. The beta coefficients are described as follows:

β_{0j} = Mean science achievement for students in school j.
β_{1j} = The degree to which sex differences in science achievement related to their achievement.
β_{2j} = The degree to which attitude towards science differences among students in school j is related to their science achievement.
β_{3j} = The degree to which ethnic differences among students in school j is related to their science achievement.
β_{4j} = The degree to which differences in verbal abilities among students in school j is related to their science achievement.
β_{5j} = The degree to which differences in quantitative abilities among students in school j is related to their science achievement.
β_{6j} = The degree to which socio-educational differences among students in school j is related to their science achievement.

The between-schools variance was reduced markedly by 76.1 percent when school effects were added to this model. In Table 8, school effects were found to be statistically significant in influencing student differences in science achievement including percentage of female teachers in the school (negative small effect), teacher decision making (negative small effect), average socio-educational level of students (large positive effect), average verbal ability (large positive effect), average quantitative ability (very large positive effect) and average ethnicity of students (small positive effect). In contrast, there was evidence that some of the school effects investigated which were not useful in explaining student differences in science achievement, such as the size of the population in the area of the school, existence of parent teacher committees, operating costs, average student attitude towards science, school type and sex composition of the school. Student differences in science achievement were explained by the following school effects model and found to fit well, that is, these school level variables were associated with greatly reduced school level unexplained variance and were statistically significant in accounting for student differences in science achievement.

$$\beta_{0j} = \gamma_{00} + \gamma_{10}Femtch_j + \gamma_{20}Tchdec_j + \gamma_{30}Avsel_j + \gamma_{40}Averb_j + \gamma_{50}Avmath_j + \gamma_{60}Minority_j + U_{0j}$$

$$\beta_{1j} = \gamma_{10} + U_{1j}$$

$$\beta_{2j} = \gamma_{20} + U_{2j}$$

$$\beta_{3j} = \gamma_{30} + U_{3j}$$

$$\beta_{4j} = \gamma_{40} + U_{4j}$$

$$\beta_{5j} = \gamma_{50} + U_{5j}$$

$$\beta_{6j} = \gamma_{60} + U_{6j}$$

These student and school effects were statistically significant in explaining student differences in science achievement and a model of school effectiveness incorporating these effects is provided in Figure 3. It should be emphasised that this study is preliminary and further effects are being studied for incorporation into the model. Of particular interest are the teacher characteristics and other process school effects.

The relationship between these effects and sex differences was examined by freeing up the variable sex (allowing the student variable sex to vary between schools). However, there were no statistically significant variations in sex differences in science achievement between-schools. Sex differences were found to contribute relatively little towards explaining differences in science achievement, with more significant effects related to verbal and quantitative abilities and social class. Moreover, student characteristics other than sex explained 41.7 percent of the student differences, while school level differences accounted for 15.2 percent. The higher male science achievement (t = -11.02, p = 0.000) was significant, but offset by other effects to a greater extent.

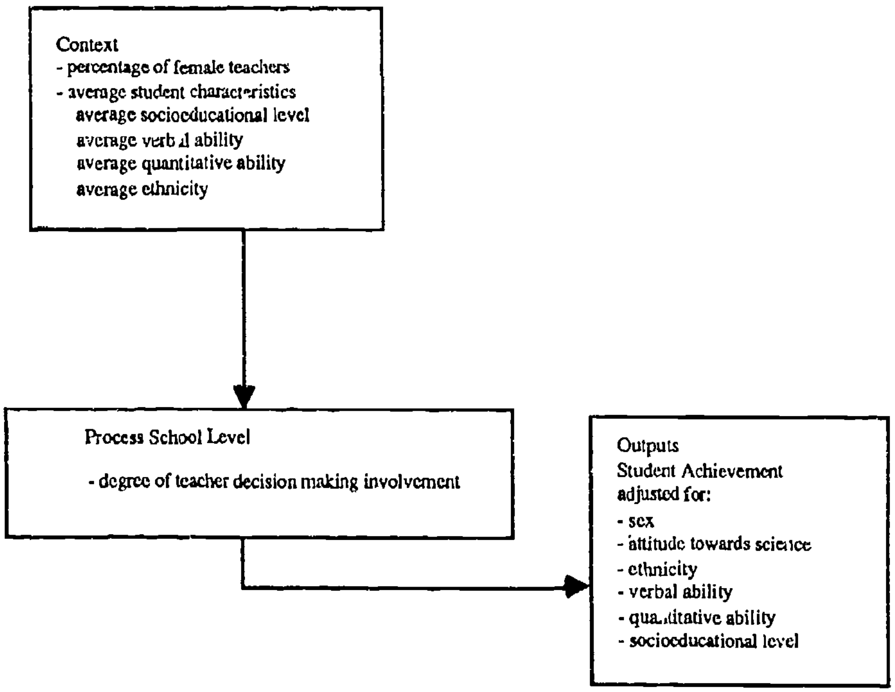

Figure 3. Preliminary Integrated Model of School Effectiveness

Importance of this Study

This analysis of the Australian student database, the Second International Science Study (SISS), established the relative importance of school-level and student-level variables in explaining student differences in science achievement. A combination of Rasch analysis and the hierarchical linear model (HLM) was used to demonstrate the relationship between contextual and process school factors and student achievement and how they are inextricably woven into the student's own characteristics. While some school effects were more useful than others in explaining student differences in science achievement, the significance of the aggregated student level variables cannot be ignored. While most of the unexplained variance was at the student-level, student variables reduced this by half. Student characteristics found to be most significant were sex of the student, attitude towards science, ethnicity, verbal ability, quantitative ability and socio-educational level. Similarly, the smaller portion of school-level unexplained variance was significantly reduced by the school-level variables. School-level variables found to be most significant were percentage of female teachers (-ve), teacher decision making (-ve), average ability (+ve), average attitude (+ve), average ethnicity (+ve) and average socio-educational level (+ve). The peer effect appeared to be greatest for average student ability. In summary, while specific school and teacher characteristics were notably influential, they appeared relatively weak in comparison to the student's own personal characteristics and the characteristics of the student's own peers. Additionally, ability measures consistently showed high correlations with science achievement, as did the socio-educational level of the student.

Although the present study is basically a preliminary one, further path analysis and use of the hierarchical linear model for all countries participating in the SISS should provide useful information regarding the relationship between student achievement, attitudes towards science (as attitude is used as an outcome variable) and other home and school environment characteristics.

References

Adams, R. J., & Toon, K. S. (1991). *Titan: The interactive test analysis system.* Melbourne: Australian Council for Educational Research.

Bryk, A.S., & Raudenbush, S.W. (1987). Application of hierarchical linear models to assessing change. *Psychological Bulletin, 101*(1), 147-158.

Bryk, A.S., & Raudenbush, S.W. (1989). Toward a more appropriate conceptualization of research on school effects: A three-level hierarchical linear model. In R. D. Bock (Ed.), *Multilevel analysis of educational data.* San Diego: Academic Press.

Bryk, A.S., Raudenbush, S.W., Seltzer, M., & Congdon, R.T. (1989, 1988, 1986). *An introduction to HLM: Computer program and users' guide.* Chicago: University of Chicago. (Available from Scientific Software in the USA)

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

Coleman, J. S., Campbell, E. Q., Hobson, C.J., McPartland, J., Mood, A.M., Weinfeld, F.D., & York, R.L. (1966). *Equality of educational opportunity.* Washington, D.C.: U.S. Government Printing Office.

Comber, L.C., & Keeves, J.P. (1973). *Science education in nineteen countries: An empirical study.* International Studies in Evaluation I. Stockholm: Almqvist and Wiksell.

Cronbach, L.J. (1976). *Research on classrooms and schools: Formulation of questions, design, and analysis* (Occasional Paper). Stanford, California: Stanford Evaluation Consortium.

Haertel, E. H., James, T., & Levin, H. M. (1987). *Comparing public and private schools Volume 2: School achievement.* New York: The Falmer Press.

Hauser, R.M., Sewell, W.H., & Alwin, D.F. (1976). High school effects on achievement. In W.H. Sewell, R.M. Hauser, & Featherman, D.L. (Eds.), *Schooling and achievement in American society* (pp. 309-341). New York: Academic Press.

Husén, T. (1967). *International study of achievement in mathematics: A comparison of twelve countries.* Stockholm: Almqvist and Wiksell.

Jencks, C., Bartlett, S., Corcoran, M., Crouse, J., Eaglesfield, D., Jackson, G., McClelland, K., Mueser, P., Olneck, M., Schwartz, J., Ward, S., & Williams, J. (1979). *Who gets ahead?: The determinants of economic success in America.* New York: Basic Books.

Jencks, C., Smith, M., Acland, H., Bane, M., Cohen, D., Gintis, H., Heyns, B., & Michelson, S. (1972). *Inequality: A reassessment of the effect of family and schooling in America.* London: Allen Lane.

Keeves, J.P. (1973). Differences between the sexes in mathematics and science courses. *International Review of Education, 19*(1), 47-63.

Kelly, A. (1978). *Girls and science: An international study of sex differences in school science achievement.* (International Association for the Evaluation of Educational Achievement IEA Monograph Studies No. 9). Stockholm, Sweden: Almqvist and Wicksell International.

Lee, V.E. (1986). *Multi-level causal models for social class and achievement*. Paper presented at the annual meeting of the American Educational Research Association, San Francisco.

Lee, V. E., & Bryk, A. S. (1989). A multilevel model of the social distribution of high school achievement. *Sociology of Education, 62*(2), 246.

Mortimore, P., Sammons, P., Stoll, L., Lewis, D., & Ecob, R. (1988). *School matters: The junior years*. Somerset, England: Open Books.

Peaker, G.F. (1967). The regression analyses of the national survey. In United Kingdom, Department of Education and Science, *Children and their primary schools* (Plowden, Chairman). London: HMSO. Volume 2, Appendix 4.

Postlethwaite, T.N., & Wiley, D.E. (1992). *The IEA study of science II: Science achievement in twenty-three countries*. Oxford: Pergamon Press.

Rasch, G. (1960, 1980). *Probabilistic models for some inteslligence and attainment tests*. Copenhagen: Danish Institute for Educational Research. Expanded edition (1980), Chicago: The University of Chicago Press.

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

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

Scheerens, J. (1990). School effectiveness research and the development of process indicators of school functioning. *School Effectiveness and School Improvement, 1*(1), 61-80.

Scheerens, J., & Creemers, B.P.M. (1989). Conceptualizing school effectiveness. *International Journal of Educational Research, 13*(7), 691-706.

Thurston, L. L. (1925). A method of scaling psychological and educational tests. *Journal of Educational Psychology, 16*, 433-451.

Thurston, L. L. (1926). The scoring of individual performance. *Journal of Educational Psychology, 17*, 446-457.

Thurston, L. L. (1927). A law of comparitive judgement. *Pyschological Review, 34*, 278-286.

Thurston, L. L. (1959). *The measurement of values*. Chicago: The University of Chicago Press.

Correspondence:

Dr Deidra Young, Research Fellow, Science and Mathematics Education Centre, Curtin University of Technology, GPO Box U1987, Perth, Western Australia, 6001, Australia.

Table 1. Description of Student and School Level Variables

Student Level Variables		
Science	Science achievement score calibrated for student ability and item difficulty using the Rasch model - 46 multiple-choice items	
Sex	Sex of student	1 = Male 2 = Female
Enjsci	Attitude towards science scale - 16 items	1 = agree 0 = disagree or not sure
Ethnty	Ethnicity composite scale - home language	1 = English 0 = not English
	- student's country of birth - father's country of birth - mother's country of birth	1 = Australia, UK, Ireland, USA, Canada, New Zealand 0 = Europe, Asia, South America, Central America, Africa, Oceania
Tot2v	Verbal ability	
Tot2q	Quantitative ability	
Sel	Socio-educational level composite scale - father's education & occupation - mother's education & occupation - number of books in the home	
School Level Variables		
Schtyp	School type - government or Catholic/independent	1 = Private (I, C) 0 = Public (G)
Sexcomp	Sex composition of the school	1 = Single-sex 0 = Coeducational
Areapopn	Approximate population of the area where your school is located	
Stutchr	Ratio of number of students to number of teachers in school	
Femtch	Percentage of female secondary teachers in school	
Tchdec	Teachers responsible for decision making in the school composite scale - range of science subjects - content of science courses - choosing textbooks for science students - selecting science equipment - selecting students for entrance to the school - determining size of tuition fees - deciding on major expenditure (e.g. new buildings) - determining conditions of employment for teachers - selecting teachers for the school - selecting the school principal - making rules and regulations for the students	1 = Teachers responsible 0 = Teachers not responsible

Ptc	Involvement in parent teacher committee (ptc) activities - school has one or more committees on which both parents and teachers are formally represented - social work in the community - social and cultural activities - activities for raising money for the school - discussions about particular curriculum and instructional matters - discussions about general policies for school organization or curriculum - organization of parent education activities - selection of teaching staff - decisions about school expenditure - organization of science-related activities	1 = ptc involved 0 = ptc not involved
Opcost	Source of operating costs for the school	1 = government (public) 2 = partially government 3 = private
Avsel	Average socio-educational level of students attending the school - aggregated variable from student Sel	
Averb	Average verbal ability of students attending the school - aggregated variable from student Tot2v	
Avmath	Average quantitative ability of students attending the school - aggregated variable from student Tot2q	
Minority	Average ethnic composition of students attending the school - aggregated variable from student Ethnty	
Avenjsci	Average attitude towards science of students attending the school - aggregated variable from student Enjsci	

Table 2. Sample Size and Composition.

Table 2.A. Australian Stratification: Population 2 stratum definitions (14-year-old students)

Stratum (h)	Description
14	Australian Capital Territory (ACT) - Government Schools
15	Australian Capital Territory (ACT) - Catholic Schools
16	Australian Capital Territory (ACT) - Independent Schools
24	New South Wales (NSW) - Government Schools
25	New South Wales (NSW) - Catholic Schools
26	New South Wales (NSW) - Independent Schools
34	Victoria (Vic) - Government Schools
35	Victoria (Vic) - Catholic Schools
36	Victoria (Vic) - Independent Schools
44	Queensland (Qld) - Government Schools
45	Queensland (Qld) - Catholic Schools
46	Queensland (Qld) - Independent Schools
54	South Australia (SA) - Government Schools
55	South Australia (SA) - Catholic Schools
56	South Australia (SA) - Independent Schools
64	Western Australia (WA) - Government Schools
65	Western Australia (WA) - Catholic Schools
66	Western Australia (WA) - Independent Schools
74	Tasmania (Tas) - Government Schools
75	Tasmania (Tas) - Catholic Schools
76	Tasmania (Tas) - Independent Schools
84	Northern Territory (NT) - Government Schools
85	Northern Territory (NT) - Catholic Schools
86	Northern Territory (NT) - Independent Schools

Table 2.B. Sampling Statistics Summary

	Target Population		Designed Sample		Achieved Sample			Response	Stratum
Stratum h	Schools A(h)	Students N(h)	Schools a(h)	Students n(h)	Schools a'(h)	Students n'(h)	f'(h)	Schools %	Weight w'(h)
14	16	2609	10	240	6	127	0.049	60	0.4104
15	6	889	3	72	2	46	0.052	67	0.3861
16	3	245	2	48	2	39	0.159	100	0.1255
24	421	65453	33	792	28	568	0.009	85	2.3020
25	158	14505	6	144	3	65	0.004	50	4.4580
26	55	3607	8	192	7	155	0.043	88	0.4649
34	410	49544	28	672	23	475	0.010	82	2.0837
35	131	11694	8	192	6	130	0.011	75	1.7970
36	80	5886	7	168	7	164	0.028	100	0.7170
44	198	28998	30	720	28	560	0.019	93	1.0345
45	86	6724	7	168	7	157	0.023	100	0.8556
46	32	2737	6	144	6	137	0.050	100	0.3991
54	143	19439	34	816	30	605	0.031	88	0.6419
55	77	2239	6	144	3	69	0.031	50	0.6482
56	18	1330	4	96	4	88	0.066	100	0.3019
64	129	17308	32	768	28	606	0.035	88	0.5706
65	46	3069	6	144	5	115	0.037	83	0.5331
66	30	1466	4	96	3	70	0.048	75	0.4184
74	70	5725	22	528	17	368	0.064	77	0.3108
75	11	850	4	96	4	86	0.101	100	0.1974
76	11	445	3	72	3	55	0.124	100	0.1616
84	10	1132	10	240	9	188	0.166	90	0.1203
85	2	212	2	48	1	22	0.104	50	0.1925
86	1	26	1	24	1	22	0.846	100	0.0236
Total	2144	246132	276	6624	233	4917	0.020	84	-

n' = 4917
N = 246132

Table 3. Construction of Science Achievement Estimate using the Rasch Model and the Computer Software Package: TITAN*

Item Estimates (Thresholds) In Input Order all on Science (N = 4917 L = 46)

Item Name	Score	Maximum Score	Threshold 1	Infit Meansq	Outfit Meansq	Infit t	Outfit t
P2m8	3356	4801	-.26	1.02	1.02	1.3	.7
P2m9	3550	4878	-.44	1.00	1.00	.3	0.0
P2m10	2970	4779	.14	1.00	1.00	.1	.1
P2m11	1024	4773	2.30	1.09	1.25	4.0	6.0
P2m12	1736	4790	1.43	1.02	1.06	1.6	2.7
P2m14	3307	4886	-.15	1.00	1.02	.3	.7
P2m15	4089	4886	-1.19	.92	.87	-3.2	-2.8
P2m16	3545	4858	-.45	.93	.93	-4.1	-2.1
P2m18	2247	4848	.93	.96	.96	-3.5	-2.2
P2m19	3874	4854	-.89	.95	.91	-2.4	-2.1
P2m20	1693	4367	1.30	.96	.96	-3.0	-1.6
P2m21	4252	4908	-1.46	.92	.92	-2.8	-1.4
P2m22	3943	4901	-.94	1.04	1.14	2.0	3.1
P2m24	1452	4739	1.74	.99	1.08	-.6	2.9
P2m25	1839	4675	1.29	1.05	1.11	3.2	5.0
P2m26	3683	4897	-.58	1.06	1.19	3.3	5.1
P2m27	3768	4899	-.69	.97	.92	-1.6	-2.3
P2m28	4247	4901	-1.46	1.02	1.08	.6	1.4
P2m29	2208	4663	.87	.97	.97	-2.2	-1.5
P2m30	2672	4816	.48	1.04	1.05	3.0	2.3
P2a1	1875	2459	-.69	.93	.88	-2.8	-2.4
P2a2	1442	2443	.26	1.00	1.01	.1	.3
P2a3	1313	2459	.55	1.02	1.05	1.2	1.9
P2a4	2285	2470	-2.24	.93	.77	-1.1	-2.0
P2a7	1893	2465	-.72	1.09	1.25	3.4	4.4
P2a9	1610	2453	-.07	1.04	1.03	1.9	.9
P2a10	1240	2382	.63	1.01	1.03	.6	1.0
P2b1	1674	2391	-.25	1.08	1.12	3.6	2.8
P2b2	1799	2420	-.50	1.00	1.07	0.0	1.5
P2b3	803	2367	1.57	.94	1.00	-2.7	.1
P2b4	2131	2448	-1.46	.99	1.29	-.2	3.2
P2b5	1998	2451	-.99	1.00	1.05	0.0	.8
P2b7	2152	2454	-1.53	1.03	1.20	.6	2.2
P2b10	1798	2447	-.46	.97	.96	-1.1	-.8
P2c3	1160	2344	.78	.96	.96	-2.3	-1.4
P2c8	1230	2434	.72	1.03	1.05	1.6	1.9
P2c9	1722	2431	-.33	.94	.89	-2.8	-2.5
P2c10	1757	2437	-.40	.94	.88	-2.5	-2.6
P2d1	1531	2432	.12	1.01	1.01	.4	.2
P2d2	1455	2434	.29	.96	.92	-2.1	-2.7
P2d3	1386	2410	.41	1.02	1.06	.9	1.9
P2d4	1576	2435	.03	.96	.93	-1.9	-2.1
P2d7	1028	2425	1.14	1.07	1.11	3.6	3.8
P2d8	931	2433	1.36	1.00	1.00	-.2	0.0
P2d9	1822	2429	-.56	.95	.90	-1.8	-1.9
P2d10	1418	2428	.35	.95	.96	-2.9	-1.5
Mean			0.00	.99	1.02	-.2	.5
SD			1.00	.05	.11	2.2	2.4

* Titan is a Rasch item analysis computer software program developed by Dr Raymond Adams and Khoo Siek Toon at the Australian Council for Educational Research and generously supplied to the authors for this study.

18

Table 4. Summary of Rasch Analysis of Science Test Items

Total Test Items

Summary of Item Estimates	
Mean	0.00
SD	1.00
SD (adjusted)	1.00
Reliability of estimate	1.00
Separation Index	21.86

Fit Statistics

Infit Mean Square		Outfit Mean Square	
Mean	.99	Mean	1.02
SD	.05	SD	.11
Infit t		**Outfit t**	
Mean	-.22	Mean	.45
SD	2.24	SD	2.41

Total Students

Summary of Case Estimates	
Mean	0.78
SD	1.09
SD (adjusted)	0.99
Reliability of estimate	0.82
Separation Index	2.15

Fit Statistics

Infit Mean Square		Outfit Mean Square	
Mean	1.00	Mean	1.02
SD	.17	SD	.39
Infit t		**Outfit t**	
Mean	.01	Mean	.05
SD	0.88	SD	.94

Table 5. Graph of Student Ability Estimates of Science Test Items

Item Estimates (Thresholds) on Science (N = 4917 students L = 46 items)

Ability Estimate	Number of Students	Items				
4.0	X					
	XX					
	X					
	XXX					
3.0						
	XXXX					
	XXXX					
	XXXX					
	XXXX	11				
	XXXXXXXX					
2.0	XXXXXXX					
	XXXXXXX					
	XXXXXXXXXXXXXXX	24				
	XXXXXXXXXXXXX	43				
	XXXXXXXXXX	12	68			
	XXXXXXXXXXXXX	20	25			
1.0	XXXXXXXXXXXXXXXX	67				
	XXXXXXXXXXXXXXXXXXXX	18	29			
	XXXXXXXXXXXXXXXXX	53	58			
	XXXXXXXXXXXXXXXXXX	33	40			
	XXXXXXXXXXXXXXXX	30	63	70		
	XXXXXXXXXXXXXXXXXX	32	62			
	XXXXXXXXXXXXXXXX	10	61	64		
.0	XXXXXXXXXXXXXX	14	39			
	XXXXXXXXXXXX	8	41			
	XXXXXXXXXXX	9	16	50	59	60
	XXXXXXX	26	42	69		
	XXXXXXX	27	31	37		
	XXXXX	19	22			
-1.0	XXX	45				
	XX	15				
	XX	21	28	44		
	X	47				
	X					
-2.0						
		34				
-3.0						

Each X represents 17 students

20

Table 6. Means, Standard Deviations, and Correlations of Student-Level Variables for the 14-year-old Students

Means and Standard Deviations

	Science Achievement	Sex	Attitude towards Science	Ethnicity	Socio-educational Level	Verbal Ability	Quantitative Ability
Mean	0.78	0.52	0.50	0.89	0.04	25.7	13.6
Standard Deviations	1.09	0.50	0.22	0.24	0.49	5.8	3.9
Min/Max Scores	-3.75/4.06	0/1	0/1	0/1	-1.00/1.34	0/40	0/20

Pearson Product Moment Correlations

	Science Achievement	Sex	Attitude towards Science	Ethnicity	Socio-educational Level	Verbal Ability	Quantitative Ability
Science Achievement	1.00						
Sex	0.13*	1.00					
Attitude towards Science	0.22*	0.15*	1.00				
Ethnicity	0.12*	-0.01	-0.07*	1.00			
Socio-educational Level	0.37*	0.00	0.13*	0.11*	1.00		
Verbal Ability	0.49*	-0.01	0.16*	0.11*	0.32*	1.00	
Quantitative Ability	0.61*	-0.03*	0.12*	0.06*	0.33*	0.46*	1.00

Table 7. HLM Analysis of Variance in Science Achievement

Model	Science Achievement: Between-Schools Variance	Science Achievement: Within- Schools Variance
Science Achievement	0.179* (15.2%)	0.998 (84.8%)
Science Achievement + Sex	0.180*	0.981 (1.7% reduction)
Science Achievement + student level† + school level‡ variables	0.045* (76.1% reduction)	0.565 (43.4% reduction)

* Statistically significant variables at $p = 0.001$ level

† Student level variables = sex, enjsci, ethnty, tot2v, tot2q, sel

‡ School level variables = areapopn, stutchr, femtch, tchdec, ptc, opcost, avsel, averb, avmath, minority, avenjsci, schtyp, sexcomp

2ɔ

Table 8. HLM Analysis of School Effects on Science Achievement

The Gamma(*)-Standard Error-T Statistic Table

	Gamma(*)	Standard Error	T Statistic	P-Value
For Base Coef.				
Base	-2.38	.35	-6.80	.000
Areapopn	-.02	.02	-1.24	.216
Stutchr	.00	.00	.04	.969
Femtch	.00	.00	-2.27	.023
Tchdec	-.28	.16	-1.73	.083
Ptc	.04	.07	.55	.580
Opcost	.16	.13	1.30	.195
Avsel	.27	.10	2.73	.007
Averb	.04	.01	3.50	.001
Avmath	.14	.02	9.09	.000
Minority	.33	.19	1.76	.079
Avenjsci	.29	.30	.98	.327
Schtyp	-.12	.13	-.92	.361
Sexcomp	.04	.06	.64	.519
For Sex Slope*				
Base	-.29	.03	-11.02	.000
For Enjsci Slope*				
Base	.54	.05	9.91	.000
For Ethnty Slope*				
Base	.16	.06	2.57	.010
For Tot2v Slope*				
Base	.05	.00	18.41	.000
For Tot2q Slope*				
Base	.12	.00	34.84	.000
For Sel Slope*				
Base	.29	.03	9.46	.000

* - The Residual Variance For This Parameter Has Been Set To Zero.

Chi Square Table

Estimated Parameter	Parameter Variance	Degrees Of Freedom	Chi Square	P-Value
Base Coef	.04497	219	559.77	.000

Sigma squared = 0.565
Parameter Reliability Estimates:
Base Tau(1,1) / D-bar(1,1) = 0.597