Hierarchical Linear Modelling of Student and School Effects on Academic Achievement

Xin Ma
Don A. Klinger

Hierarchical linear modelling (HLM) and data from the New Brunswick School Climate Study were used to examine student background, school context, and school climate effects on Grade 6 student achievement in mathematics, science, reading, and writing. Gender, socioeconomic status (SES), and Native ethnicity were significant predictors of academic achievement. Schools showed the smallest variation in reading, the largest in mathematics. School mean SES was significant in mathematics, reading, and writing achievement, as was disciplinary climate in mathematics, science, and writing. School size and parental involvement significantly affected only the relationship between mathematics achievement and individual SES.

La modélisation linéaire hiérarchique et les données provenant d'une étude portant sur le climat scolaire au Nouveau-Brunswick *(New Brunswick School Climate Study)* furent utilisées pour analyser les acquis des élèves, le contexte scolaire et les effets du climat scolaire sur le rendement d'élèves de 6e année en mathématique, en sciences, en lecture et en écriture. Le sexe, la situation socio-économique (SSE) et l'origine autochtone étaient des prédicteurs importants du rendement scolaire. Les écoles affichent les plus faibles variations en lecture et les plus fortes, en mathématique. La SSE moyenne de l'école était un facteur important dans le rendement en mathématique, en lecture et en écriture, tout comme le climat disciplinaire pour la mathématique, les sciences et l'écriture. La taille de l'école et la participation des parents n'avaient une incidence importante que sur le rapport entre le rendement en mathématique et la SSE personnelle.

This study examined the influence of student and school factors on Grade 6 students' performance in mathematics, science, reading, and writing in New Brunswick. Student characteristics and school context and climate factors were included in an effort to fill gaps in the research on effective elementary schools. Increasing emphasis on academic performance (Council of Ministers of Education, Canada, 1998; Educational Testing Service, 1999) gives new importance to investigating factors that contribute to student learning. Because education systems have a hierarchical structure (students nested within schools), researchers must examine both student and school characteristics (Bryk & Raudenbush, 1992). Student

characteristics can be individual characteristics, such as gender, or family characteristics, such as socioeconomic status (SES). Gender differences in achievement tend to be subject specific (Hedges & Nowell, 1995; Manning, 1998; Sammons, West, & Hind, 1997). Males outperform females in mathematics and science, with larger differences in science (Beller & Gafni, 1996); females outperform males in reading and writing, with larger differences in writing (Battistich, Solomon, Kim, Watson, & Schaps, 1995; Sammons et al., 1997). In general, females exhibit superior verbal skills and males display superior spatial and quantitative skills (Hedges & Nowell, 1995).

SES has long been used to explain differences in student academic achievement (Sammons et al., 1997; Thomas, Sammons, Mortimore, & Smees, 1997). In 1982, White described the effect as the most enduring finding in sociological research. Lytton and Pyryt (1998) have shown that in Canada between 35% and 50% of the variation in elementary students' academic achievement can be attributed to SES. Race and ethnicity also affect academic achievement (Sammons et al., 1997; Strand, 1997). Some researchers attribute the low academic achievement of certain racial groups to their low SES (for example, Hull's 1990 discussion of Canadian Native students); others attribute their underachievement to their unsuccessful incorporation, voluntary or involuntary, into the dominant culture (Ogbu & Simons, 1998). Family structure (number of parents and number of siblings) is also said to influence student academic achievement (Manning, 1998; Pong, 1997, 1998).

Teacher satisfaction, principal leadership, disciplinary climate, academic press (expectations), and parental involvement often constitute school climate. These factors can be controlled directly by school staff and influence student academic achievement (Coyle & Witcher, 1992; Downer, 1991; Willms, 1992; Zigarelli, 1996). In general, research on effective schools indicates three important school climate factors (Willms, 1992). Students learn more and perform better in schools that have strong parental involvement (Goldring & Shapira, 1996; Ho & Willms, 1996), emphasize academic success (Lytton & Pyryt, 1998; Zigarelli, 1996), and have a disciplinary climate conducive to teaching and learning (DeBaryshe, Patterson, & Capaldi, 1993; Ma & Willms, 1995).

Other school factors, such as size, location, and mean SES, are contextual and usually beyond the immediate control of school staff; they too influence academic achievement (Sammons et al., 1997; Willms, 1992). The average SES of a school has been found to have as great an effect as an individual student's SES (Caldas & Bankston, 1997; Ho & Willms, 1996): Students attending a school with a higher mean SES are more likely to succeed academically, and this effect is over and above that of individual

student SES (Willms, 1992). School size has not shown a consistent effect on academic achievement (Griffiths, 1996; Luyten, 1994).

One problem with much previous research is its inability to accommodate the hierarchical or nested structure of educational data. Actions and measures at one level affect, and are affected by, actions and measures at other levels, and this interaction must be considered in data analysis (Bryk & Raudenbush, 1992). Research also needs to examine school climate in relation to student academic achievement (Willms & Raudenbush, 1989). Most previous studies have focused primarily on student characteristics, school-context characteristics, or a single school climate factor and so do not provide adequate knowledge about how school policies and practices influence student learning outcomes. Finally, most previous studies of effective schools focused on secondary schools, but educational problems, especially academic problems, are cumulative. Students in effective elementary schools are better prepared for secondary school.

Some recent national and provincial surveys of education do provide richer descriptions of school climate than earlier ones. And recently developed statistical techniques such as hierarchical linear modelling (HLM) allow researchers to analyze multilevel data such as students nested within schools.

DATA

About one-third of New Brunswick's population speaks French. The province is officially bilingual, and there are separate English and French school systems. Other than the Mi'kmaq and Maliseet Native communities, there are few visible minorities. The entire Grade 6 student population in the English system participated in the New Brunswick School Climate Study (NBSCS), conducted during the 1995–96 school year (N = 6,883 students from 148 schools). Each student completed achievement tests in mathematics, science, reading, and writing as well as a questionnaire. We used the student achievement scores as dependent measures and the student characteristics and school context and climate items in the student questionnaire as independent measures.

METHOD

Measures and Variables

A panel of experienced teachers and subject-area specialists developed the four NBSCS achievement tests on the basis of the provincial curriculum. The mathematics test consisted of 39 items designed to measure

numeration, measurement, geometric ability, and data management, with an emphasis on understanding concepts and solving problems. Students were encouraged to use manipulatives, and calculators were permitted. Cronbach's alpha was 0.86. The science test consisted of 33 items designed to measure knowledge and understanding of scientific concepts and processes. Cronbach's alpha was 0.79. The reading test used 35 items in fictional and non-fictional passages to measure comprehension. Cronbach's alpha was 0.84. The writing assessment was based on two pieces of student writing, one chosen by the teacher from regular class work, the other written during the assessment. A panel of teachers used scoring rubrics and a 6-point scale to grade them from *unrateable* to *superior*. Each student's final writing score was the sum of the two scores, scaled to have a mean of 0 and a standard deviation of 1 following the statistical procedures outlined in Mosteller and Tukey (1977).

The questionnaire consisted of 22 questions, most containing embedded items, to measure student, family, and school characteristics. We used the following to obtain information on major student variables:

Socioeconomic status. Which of these things do you have at home for you to use? Books of my own, my own magazine(s), a dictionary, a computer, a calculator, a musical instrument, a phone, a specific place to study, and a link to the Internet. Which of these activities have you done with members of your family over the past year? Visited parks together, gone shopping, gone to the public library, attended music concerts, gone skiing, gone on a Canadian holiday outside of New Brunswick, and gone on a holiday outside of North America.

Ethnicity. The only ethnic group identified was students from the 11 Native communities in New Brunswick, such as Burnt Church and St. Mary's.

Number of parents. Most of the time I live with: my mother, my step-mother, or someone who is like a mother to me; my father, my step-father, or someone who is like a father to me.

Number of siblings. I have (number) sisters and (number) brothers (including step- and half-sisters and brothers if any).

Gender, Native ethnicity, and the number of parents were coded as dichotomous variables. SES was estimated using students' reports of education-related possessions at home and their participation in social and cultural activities, rather than parental income or occupation. The number of siblings was a continuous variable.

School size was based on enrollment in Grade 6, and school mean SES was derived from the SES of individual students. Finally, three school

climate variables — disciplinary climate, academic press, and parental involvement — were constructed by averaging scores on 5-point scales for selected items about the school:

Disciplinary climate. Rules at this school seem to be always changing, students at this school call each other names, students fool around during class, children at our school know what "good behavior" means, students behave well in class, students at this school get into fights, rules at this school are fair, troublemakers disrupt my teacher's lessons, often the punishment for breaking the rules is too strict, the rules for behavior at this school are clear to me, children know what will happen if they break a rule, students are able to help make the rules here, students agree with the rules at this school, and kids in this school bully others outside the classroom.

Academic press. How many students in your class: Think it is important to do well at school, try hard to get good marks on tests, could do better if they tried harder, find school work too difficult, usually do their homework on time, think it is more important to have fun than learn, and feel they can do the work in class if they try? Our teacher expects all students to do well, most of my school work is too easy for me, we often have lively discussions in class, the teacher encourages students to try harder, school work is challenging for me, doing my homework helps me learn what we are taught in class, the teacher encourages us to ask questions about the material we are studying, and I can do well in school if I work hard.

Parental involvement. Since the beginning of school this past fall, how often have your parents (or guardians) done the following things? "Helped" you with your homework, talked with you about how students treat you, limited how much TV you could watch weekdays, discussed how well you were doing in math, discussed how well you were doing in reading, talked with you about school projects, checked your homework for mistakes, said how important school work is, discussed hurtful things that children might say, helped in the classroom, and helped with school activities (e.g., field trip).

So constructed, disciplinary climate concerns mainly rules and compliance. Cronbach's alpha was 0.77 for disciplinary climate, 0.61 for academic press, and 0.77 for parental involvement.

Effect Size

The effects of student-level and school-level variables may be represented in various ways. We chose *effect size* to make it easier for us to compare the effects of explanatory variables on the outcome measure (cross-variable comparison) because explanatory variables were converted to the same scale and for others to compare their results with ours (cross-study comparison). We standardized outcome and explanatory variables (converting

raw scores into z-scores for each) so that they had a mean of 0 and a standard deviation of 1. After fitting a regression model, coefficients associated with explanatory variables were measures of effect size.

Statistical Procedure

Educational data are often hierarchical or multilevel (students nested within schools). Failure to consider their hierarchical nature leads to unreliable estimation of the effectiveness of school policies and practices (Bryk & Raudenbush, 1992; Raudenbush & Willms, 1991). Unfortunately, most analyses in educational research have not taken into account the hierarchical structure of educational data.

Over the past decade, researchers have developed hierarchical linear models, multilevel models that can simultaneously estimate the effects of student-level and school-level variables (Bryk & Raudenbush, 1992). A separate regression model is fitted for each school. These regression models yield a mean score adjusted for student background for each school. They also produce measures of equality: for example, of males and females or of academic achievement in relation to social class. Individual school estimates (adjusted mean scores or measures of equality) then become dependent measures in a model that attempts to explain variation among schools with measures of school characteristics (Gamoran, 1991; Lee & Smith, 1993; Willms, 1992).

We used two-level HLM (Bryk & Raudenbush, 1992) to examine the effects of student and school variables on academic achievement at the student and school (students nested within schools) levels. Separate HLM analyses were conducted for the achievement measure in each of mathematics, science, reading, and writing. Each HLM analysis was done in three stages. At the first stage, the analysis produced the null model with no independent variables at the student and school levels. With only the student-level outcome measure, this model was similar to a random-effect ANOVA model, providing a measure of the variances within and between schools for each of the four achievement measures.

At the second stage, student variables were added to the null model, first separately, to determine whether each variable had a significant absolute effect on academic achievement measures independently of other variables and whether its relationship with achievement varied significantly across schools, then in combination, to determine whether each variable had a significant relative effect on the academic achievement measures in the presence of other variables. In other words, the relative effect of the variable was adjusted for the shared effects of other variables. We examined

these specific effects to find the role of each variable and the interrelationship between it and others.

Using a similar procedure, at the third stage of the analysis school variables were added to the student model, first separately, to examine their absolute effects, then in combination, to examine their relative effects — that is, to model average school academic achievement measures and school variables, and relationships between academic achievement measures and student variables in relation to school variables.

RESULTS

Table 1 shows the means and standard deviations for the outcome (dependent) variables and explanatory (independent) variables as well as coding information for the dichotomous variables (gender, Native ethnicity, and number of parents). Means for the dichotomous variables are proportions of the category coded as 1. Descriptive statistics are in their original scales; writing achievement, SES, and school mean SES were available only as z-scores, with a mean of 0 and a standard deviation of 1. All continuous variables, with the exception of school size, were standardized to have a mean of 0 and a standard deviation of 1 at both the student level and the school level. School size was determined by the number of students in Grade 6, and we used its original scale. Dichotomous variables were not standardized.

Most of the variation in achievement was among students within schools: 0.89 in mathematics, 0.91 in science, 0.95 in reading, and 0.91 in writing. However, schools differed markedly by subject for the balance of the variation. The smallest variation (0.05) suggests that schools were not very different in reading achievement, the largest (0.11) suggests large differences among students in mathematics achievement.

The relative and absolute effects for student and school variables are shown in Table 2. Effect size is the amount of change in academic achievement, expressed as a proportion of a standard deviation, associated with one standard deviation increase in an explanatory variable.

Absolute effects provided a measure of the independent effects of student variables. Each student variable had a significant absolute effect on academic achievement across subject areas. Gender differences varied greatly in absolute effects across subject areas, favouring males in mathematics and science, and favouring females in reading and writing. Males scored 5% of a standard deviation higher than females in mathematics achievement and 13% of a standard deviation higher than females in science achievement. To understand the magnitude of these effect sizes,

TABLE 1

Means and Standard Deviations of Outcome and Explanatory Variables

Variable	*M*	*SD*
Academic achievement (Outcome variables)		
Mathematics	17.88	5.87
Science	17.92	5.39
Reading	23.52	6.21
Writing	0.00	1.00
Student characteristics (Explanatory variables)		
Gender (0 = male; 1 = female)	0.50	0.50
Socioeconomic status	0.00	1.00
Native ethnicity (0 = non-Native; 1 = Native)	0.01	0.10
Number of parents (0 = two parents; 1 = single parent)	0.13	0.34
Number of siblings	1.75	1.51
School characteristics (Explanatory variables)		
School size	39.71	30.73
School mean socioeconomic status	0.00	1.00
Academic press	3.72	0.16
Disciplinary climate	2.96	0.30
Parental involvement	2.27	0.17

Note. Gender, Native ethnicity, and number of parents are dichotomous variables. Socioeconomic status and writing achievement are standardized variables (i.e., they are z-scores with a mean of 0 and a standard deviation of 1) at the student level, and school mean SES is a standardized variable at the school level.

one may consider a standard achievement test, such as the Scholastic Aptitude Test (SAT), with a mean of 500 and a standard deviation of 100. If the female average was 500 in both mathematics and science, then the male average would be 505 (500 + [100 × 5%]) in mathematics and 513 (500 + [100 × 13%]) in science. Other effect sizes can be understood in the same manner. In reading and writing, males scored 19% and 42% of a standard deviation below females.

Individual SES had absolute effects ranging from 13% to 18% of a standard deviation across the four subject areas. Native students scored below non-Native students in all four subject areas: 25% of a standard deviation below in mathematics achievement, 39% in science, 36% in reading, and 34% in writing. Students from single-parent households scored below those from two-parent families: 13% of a standard deviation below in mathematics, 15% in science, 19% in reading, and 14% in writing. The absolute effects of the number of siblings were trivial.

TABLE 2

HLM Effects of Student and School Variables on Academic Achievement

	Mathematics		*Science*		*Reading*		*Writing*	
Variables	*Absolute effect*	*Relative effect*	*Absolute effect*	*Relative effect*	*Absolute effect*	*Relative effect*	*Absolute effect*	*Relative effect*
Student variables								
Gender	−0.05* (0.02)	−0.06* (0.03)	−0.13* (0.03)	−0.14* (0.02)	0.19* (0.03)	0.17* (0.02)	0.42* (0.02)	0.40* (0.02)
SES	0.14* (0.02)	−0.09* (0.02)	0.13* (0.01)	0.14* (0.01)	0.18* (0.01)	0.17* (0.01)	0.16* (0.01)	0.14* (0.01)
Native ethnicity	−0.25* (0.10)	—	−0.39* (0.12)	−0.34* (0.12)	−0.36* (0.09)	−0.32* (0.12)	−0.34* (0.11)	−0.32* (0.10)
Number of parents	−0.13* (0.04)	—	−0.15* (0.04)	—	−0.19* (0.04)	—	−0.14* (0.04)	—
Number of siblings	−0.02* (0.01)	—	−0.04* (0.01)	—	−0.04* (0.01)	—	−0.03* (0.01)	—
School variables								
School size	0.00 (0.00)	—	0.00 (0.00)	—	0.00 (0.00)	—	0.00 (0.00)	—
School mean SES	0.05 (0.03)	0.06* (0.03)	0.05 (0.03)	—	0.07* (0.03)	0.07* (0.02)	0.10* (0.03)	0.10* (0.03)
Academic press	0.07* (0.03)	—	0.09* (0.02)	—	0.03 (0.02)	—	0.09* (0.03)	—
Disciplinary climate	0.07* (0.03)	0.07* (0.03)	0.10* (0.03)	0.10* (0.03)	0.03 (0.02)	—	0.06* (0.03)	0.08* (0.02)
Parental involvement	0.01* (0.03)	—	0.01 (0.03)	—	0.02 (0.02)	—	0.05 (0.03)	—

* $p < .05$

Note. Relative effects are estimated based on the final, simplified models. Dashes indicate nonsignificant relative effects. Values in parentheses are the corresponding standard errors for the effect sizes.

Relative effects were adjusted (controlled) for other variables in the model. Gender differences remained, and gender gaps did not change much. The effects of individual SES also remained, and its relative effects were not much different from its absolute effects. So, gender and SES remain critical in explaining differences in academic achievement.

The Native gap in mathematics achievement disappeared in the presence of gender and individual SES. However, when gender and individual SES were controlled for, large Native gaps in science, reading, and writing achievement remained and were not much different from the absolute effects. The number of parents no longer had a significant effect; neither did the number of siblings. The cumulative relative effects of gender, individual SES, and Native ethnicity were estimated and found to be substantial. For example, a Native male whose SES was one standard deviation below the average could be 34% of a standard deviation below the average in science achievement, 64% in reading achievement, and 86% in writing achievement.

Average school achievement in each subject area was independent of school size and parental involvement, neither of which showed a significant absolute effect. Students in schools with higher mean SES performed significantly better in mathematics, reading and writing. And these effects were over and above those of individual SES. Disciplinary climate and academic press both had significant absolute effects in mathematics, science, and writing, and the relative effect of disciplinary climate was quite similar to its absolute effect. These effects were over and above the effects of student variables, and in the case of writing achievement, the effect of disciplinary climate was also over and above the effect of school mean SES.

The relationships (or slopes) between academic achievement and student variables such as SES may differ across schools. A shallow slope indicates that a difference in SES does not make a big difference in academic achievement, a steep slope that the difference in SES does make a big difference in academic achievement. We examined the slope between student achievement in each area and each student variable. The only significant slope was between mathematics achievement and individual SES: This relationship varied significantly across schools. When the slope was modelled with the school variables, individual SES had a greater effect on mathematics achievement in larger schools (0.01, $p < .05$) and in schools with stronger parental involvement (0.03, $p < .05$).

DISCUSSION

Most of the variables we examined at the student and school levels are complex both conceptually and operationally. To avoid simplistic implica-

tions for policies and practices that can reinforce various social stereotypes, our recommendations are tied to the measurement of the associated variables.

At the student level, we consider four findings important. First, the items used to measure SES mean that this variable was about neither income nor parental occupation but rather education-related possessions and participation in social-cultural activities. This measure of SES may underestimate the effect of SES at the student level but can still generate a reliable estimate of its effect at the school level (Willms, 1992). The items emphasized affective elements — that is, families' attitudes and beliefs about schooling and learning. Low student academic achievement correlated with negative family attitudes and beliefs. This finding suggests opportunities to work with parents and students to improve student academic achievement. It also illustrates the importance of considering the social construction of SES: Relating student academic achievement to family income or parental occupation would lead to a totally different set of remedial measures.

Second, the effects of family structure (number of parents and number of siblings) disappeared when SES was considered (cf. Ma, 1997; Manning, 1998; Sammons et al., 1997). Third, the research literature in general claims that racial-ethnic differences in academic achievement disappear once SES is considered (Hull, 1990; Sammons et al., 1997; Strand, 1997). Our study provides quite different evidence: When SES was taken into account, the relative effect of Native ethnicity on science, reading, and writing achievement remained as strong as its absolute effect. Native ethnicity was the single most important variable in this study, with more than twice the effect of SES in three out of four subject areas. Such a result has rarely been observed and suggests that the underachievement of Native students is not attributable merely to their SES but, perhaps, to their unsuccessful incorporation into the mainstream culture (Ogbu & Simons, 1998).

Fourth, gender gaps existed even after controlling for SES. However, gender differences in mathematics achievement were so small, in contrast to consistently substantial gender gaps in other subject areas, that they may indicate a trend toward gender equity in mathematics education. This hypothesis might find support in the observations that there was no Native disadvantage in mathematics achievement and that the smallest SES effect was in mathematics. Mathematics was clearly the subject area with the greatest equity in learning outcomes, perhaps the result of decades of emphasis on equity issues in mathematics education (Gambell & Hunter, 1999).

School mean SES had significant effects over and above student-level effects in reading and writing but not in mathematics and science. This result generates further research questions. It makes sense where there is

a large population of immigrants: Many immigrant children go to low-SES schools, and English is often more difficult for them than other subjects. However, New Brunswick's immigrant population is very small. Why does learning reading and writing seem more sensitive to school SES than learning mathematics and science?

There was more socioeconomic inequality among students in large schools, but school size did not affect a school's average academic achievement. Large elementary schools tend to be located in urban settings that have more socioeconomic differences. Large urban schools often offer curricular and extracurricular activities not available in small schools, and all students benefit. Yet high-SES students have more resources and can better profit from opportunities like field trips out of the province. And if large schools do not develop adequate personal connections with individual students, an impersonal environment may negatively affect low-SES students.

We found strong correlations between individual SES and student achievement in schools with strong parental involvement. The items we used to measure parental involvement assessed mainly the amount of interaction between parents and children. High-SES parents are more likely to be involved in schools and to promote their children's academic success (Stevenson & Baker, 1987). This may explain the large SES gap in academic achievement among students in schools in our study with strong parental involvement.

Of the three school climate factors we investigated, disciplinary climate, which here concerned mainly rules and compliance, was seen to be the most important determinant of academic achievement. This is an addition to the inventories of traits of effective schools reported by Lytton and Pyryt (1998) and Zigarelli (1996). Some argue that clear reasonable rules and sanctions, active and proper enforcement, and positive relationships between students and school staff form the basic elements of a disciplinary climate conducive to academic success (Ma & Willms, 1995).

Parental involvement is often emphasized, but we wonder whether frequency of school-home contacts is an unambiguous measure of parental involvement. Frequent teacher-parent communication may not be positive, since teachers often contact low-SES parents to discuss their children's learning or behavioural problems (Ho & Willms, 1996). Without knowing the nature of the communication, it is risky to take frequency as a measure of the extent of parental involvement. Norris (1999) emphasized too that the amount of parental involvement reported depends on who answers the questions. Teachers reported less involvement by parents of low-SES children, but no such differences were found when *parents* were asked about the amount of educational support they provided for their children.

Different forms of parental involvement also need to be considered. Ho and Willms (1996) documented four distinct types: home supervision, home discussion, home-school communication, and volunteer work.

CONCLUSION

We believe that HLM is an important statistical tool for investigating the relationship between student achievement and school context and climate. By taking into account the hierarchical nature of educational data, HLM separates variation in schooling outcomes into between-student and between-school components and then analyzes each component in relation to the other. Thus HLM can offer better statistical adjustments and more accurate estimations and promote better policies and practices.

ACKNOWLEDGMENTS

The authors are grateful to J. Douglas Willms, Director of the Canadian Research Institute for Social Policy at the University of New Brunswick, for providing the data for this study. The authors also thank the reviewers of this paper, whose insightful comments helped to refine the research.

REFERENCES

Battistich, V., Solomon, D., Kim, D., Watson, M., & Schaps, E. (1995). Schools as communities, poverty levels of student populations, and students' attitudes, motives, and performance: A multilevel analysis. *American Educational Research Journal, 32,* 627–658.

Beller, M., & Gafni, N. (1996). The 1991 International Assessment of Educational Progress in mathematics and sciences: The gender differences perspective. *Journal of Educational Psychology, 88,* 365–377.

Bryk, A. S., & Raudenbush, S. W. (1992). *Hierarchical linear models.* Newbury Park, CA: Sage.

Caldas, S. J., & Bankston, C., III. (1997). Effects of school population socioeconomic status on individual academic achievement. *Journal of Educational Research, 90,* 269–277.

Council of Ministers of Education, Canada. (1998). *Report on education in Canada.* Toronto: Author.

Coyle, L., & Witcher, A. (1992). Transforming the idea into action: Policies and practices to enhance school effectiveness. *Urban Education, 26,* 390–400.

DeBaryshe, B. D., Patterson, G. R., & Capaldi, D. M. (1993). A performance model for academic achievement in early adolescent boys. *Developmental Psychology, 29,* 795–804.

Downer, D. F. (1991). Review of research on effective schools. *McGill Journal of Education, 26,* 323–331.

Educational Testing Service. (1999). *Too much testing of the wrong kind, too little of the right kind in K–12 education.* Princeton, NJ: Author.

Gambell, T. J., & Hunter, D. M. (1999). Rethinking gender differences in literacy. *Canadian Journal of Education, 24,* 1–16.

Gamoran, A. (1991). Schooling and achievement: Additive versus interactive models. In S. W. Raudenbush & J. D. Willms (Eds.), *Schools, classrooms, and pupils: International studies of schooling from a multilevel perspective* (pp. 37–52). San Diego: Academic Press.

Goldring, E. B., & Shapira, R. (1996). Principals' survival with parental involvement. *School Effectiveness and School Improvement, 7,* 342–360.

Griffiths, J. (1996). Relation of parental involvement, empowerment and school traits to student academic performance. *Journal of Educational Research, 90,* 33–41.

Hedges, L. V., & Nowell, A. (1995). Sex differences in mental test scores, variability, and numbers of high scoring individuals. *Science, 269,* 41–45.

Ho, S., & Willms, J. D. (1996). Effects of parental involvement on eighth-grade achievement. *Sociology of Education, 69,* 126–141.

Hull, J. (1990). Socioeconomic status and Native education. *Canadian Journal of Native Education, 17,* 1–14.

Lee, V. E., & Smith, J. B. (1993). Effects of school restructuring on the achievement and engagement of middle-grade students. *Sociology of Education, 66,* 164–187.

Luyten, H. (1994). School size effects on achievement in secondary education: Evidence from the Netherlands, Sweden, and the USA. *School Effectiveness and School Improvement, 5,* 75–99.

Lytton, H., & Pyryt, M. (1998). Predictors of achievement in basic skills: A Canadian effective schools study. *Canadian Journal of Education, 23,* 281–301.

Ma, X. (1997). A multiple regression analysis of mathematics achievement in the Dominican Republic. *International Journal of Educational Development, 17,* 313–321.

Ma, X., & Willms, J. D. (1995, April). *The effects of school disciplinary climate on eighth grade achievement.* Paper presented at the annual meeting of the American Educational Research Association, San Francisco.

Manning, M. L. (1998). Gender differences in young adolescents' mathematics and science achievement. *Childhood Education, 74,* 168–171.

Mosteller, F., & Tukey, J. W. (1977). *Data analysis and regression.* Reading, MA: Addison-Wesley.

Norris, C. (1999). Parents and schools: The involvement, participation and expectations of parents in the education of their children. *Education Quarterly Review, 5,* 61–80.

Ogbu, J. U., & Simons, H. D. (1998). Voluntary and involuntary minorities: A cultural-ecological theory of school performance with some implications for education. *Anthropology and Education Quarterly, 29,* 155–188.

Pong, S. (1997). Family structure, school context, and eighth-grade math and reading achievement. *Journal of Marriage and the Family, 59,* 734–746.

Pong, S. (1998). The school compositional effect of single parenthood on 10th-grade achievement. *Sociology of Education, 71,* 23–42.

Raudenbush, S. W., & Willms, J. D. (Eds.). (1991). *Schools, classrooms, and pupils: International studies of schooling from a multilevel perspective.* San Diego: Academic Press.

Sammons, P., West, A., & Hind, A. (1997). Accounting for variations in pupil attainment at the end of Key Stage 1. *British Educational Research Journal, 23,* 489–511.

Stevenson, D., & Baker, D. (1987). The family-school relation and the child's school performance. *Child Development, 58,* 1348–1357.

Strand, S. (1997). Pupil progress during Key Stage 1: A value added analysis of school effects. *British Educational Research Journal, 23,* 471–487.

Thomas, S., Sammons, P., Mortimore, P., & Smees, R. (1997). Differential secondary school effectiveness: Comparing the performance of different pupil groups. *British Educational Research Journal, 23,* 451–469.

White, K. R. (1982). The relation between socioeconomic status and academic achievement. *Psychological Bulletin, 91,* 461–481.

Willms, J. D. (1992). *Monitoring school performance: A guide for educators.* London: Falmer.

Willms, J. D., & Raudenbush, S. W. (1989). A longitudinal hierarchical linear model for estimating school effects and their stability. *Journal of Educational Measurement, 26,* 209–232.

Zigarelli, M. A. (1996). An empirical test of conclusions from effective schools research. *Journal of Educational Research, 90,* 103–110.