Xin Ma
University of Kentucky

and

Robert Crocker
Memorial University of Newfoundland

Provincial Effects on Reading Achievement

Although Canadian students performed well on the 2000 Programme for International Student Assessment (PISA), there were considerable differences among provinces. Based on the Canadian sample of PISA 2000, our multilevel analysis examined the extent to which schools in various provinces produced differential effects on the reading achievement of their students. School characteristics that showed significant provincial effects on reading achievement included variables that described school context, school resources, and school climate, with most of these significant school-level variables being measures of school climate. Disciplinary climate and sense of belonging to school were the most important of significant school characteristics. They showed significant average effects by province on reading achievement, and their effects varied significantly across provinces. Results of our analysis suggest that provinces can use school effects to promote the reading achievement of their students.

Bien que les élèves canadiens aient bien réussi au Programme international pour le suivi des acquis des élèves (PISA) en 2000, les résultats accusaient d'écarts importants entre les provinces. Nous appuyant sur l'échantillon canadien du PISA 2000, nous avons analysé, à plusieurs niveaux, la mesure dans laquelle des écoles de différentes provinces ont produit des effets différentiels sur les résultats de leurs élèves en lecture. Parmi les caractéristiques de l'école qui ont joué un rôle significatif dans les résultats provinciaux d'aptitude à la lecture, notons les variables suivantes : le contexte de l'école, les ressources de l'école et le climat scolaire. L'ambiance disciplinaire et le sentiment d'appartenance face à l'école se sont avérés être les caractéristiques les plus importantes de l'école, jouant un rôle significatif dans les résultats provinciaux d'aptitude à la lecture. D'écarts importants ont été notés dans l'effet de ces deux variables d'une province à l'autre. Les résultats de notre analyse indiquent que les provinces peuvent s'appuyer sur ces variables du climat scolaire pour améliorer l'aptitude à la lecture de leurs élèves.

This study was an investigation of school factors associated with differences in reading achievement among Canadian provinces, with adjustment for student and family factors, using data from the Canadian sample of the Programme for International Student Assessment (PISA). Although Canadian students performed well on the 2000 PISA, ranking second in reading, sixth in mathematics, and fifth in science among 32 countries, there were considerable differences among provinces (Human Resources Development Canada, Statistics Canada,

Xin Ma is a professor of education and a fellow of the (US) National Academy of Education. His areas of research include mathematics education, school effectiveness, policy research, and advanced quantitative methods.
Robert Crocker is a professor of education. His areas of research include policy research, models of school achievement, and testing and evaluation.

& Council of Ministers of Education Canada, 2001). In general, students in Alberta and Quebec performed above the national average; those in Ontario, Manitoba, Saskatchewan, and British Columbia performed at about the national average; and the performance of students in the Atlantic region was below the national average.

These differences are reasonably consistent with results found on other national and international assessments such as the School Achievement Indicators Program (SAIP, Crocker, 2002a) and the Third International Mathematics and Science Study (TIMSS, Robitaille & Taylor, 2000; Robitaille, Taylor, & Orpwood, 1996, 1997). Such persistent and sizable differences raise a concern about uneven academic performance across provinces. A recent review of Canadian research on factors contributing to the learning outcomes of students has pointed to the need for a comprehensive program of research, using existing large-scale databases, to address gaps in working knowledge regarding provincial differences in academic achievement (Crocker, 2002b). This study was one contribution to this needed research.

We situate this study in the thriving research context of school effectiveness given our focus on provincial and school variables that affect reading achievement of students. Specifically, we adopt the input-process-output model of school effectiveness as our conceptual framework (Purkey & Smith, 1983; Teddlie & Reynolds, 2000). In this framework, inputs such as family characteristics, home influences, and family social and cultural values are what students bring into their schools. Schools, with differential context and climate, then channel (or process) students with varying inputs into various categories of schooling outcomes (outputs). The focus of this model on family and school is understandable because they are the most critical social institutions that influence the lives of students.

In a comprehensive examination of children's educational attainment using the British Household Panel Survey (BHPS), Croll (2004) concluded that "the central point that emerges from [his] analysis is the importance of families" (p. 412). Family plays a crucial role in children's literacy development because it is one of the best places to create "zones of proximal development" where children interact with "experts" (i.e., parents, older siblings, grandparents, and relatives) in daily routines and activities to develop literacy skills (Gallimore & Goldenberg, 1993; Rogoff, 1990; Tharp & Gallimore, 1988; Weisner, 1984). For example, dinner conversations are found to relate positively to linguistic abilities of children in low-income neighborhoods (Snow, Barnes, Chandler, Goodman, & Hemphill, 1991). The literature on family effects clearly identifies family characteristics and family educational support as key indicators of children's reading achievement.

The literature indicates that children's reading achievement is strongly related to characteristics of their families. In general, children from disadvantaged families are more likely to experience academic difficulties in reading, with low socioeconomic status (SES), minority status, and single parents as the most discussed forms of disadvantage (Entwisle & Alexander, 1988; Jencks & Philips, 1998; Podell & Soodak, 1993; Pong, 1997). The effects of family characteristics on children's educational outcomes are usually understood in terms of the transmission of social capital (social relations and networks that

provide children with opportunities to develop an identity that allows them to understand and value cognitive development, Coleman, 1990; Horvat, Weininger, & Lareau, 2003; Schneider & Stevenson, 1999). Disadvantaged families often have limited social capital to encourage, promote, or support children to appreciate learning and schooling.

Arzubiaga, Rueda, and Monzo (2002) argued that

> the daily family practices in which children participate may affect their access to school-based literacy activities; their notions of engagement and the organization of literacy practices; their appreciation of and interest in reading; and their idea of what counts as meaningful literacy. (p. 5)

This argument highlights Croll's (2004) notion that what parents do with children at home makes a major difference in their educational outcomes. The literature emphasizes home learning environment (e.g., reading materials at home), parental mentoring (e.g., monitoring reading homework), and linguistic practices of a family (e.g., reading together as a family) as having significant effects on the development of children's language capacities (Croll, 2004; Washington, 2001). Overall, "family involvement in children's learning, and a positive supportive family atmosphere, will provide a critical source of education and social support that promotes children's development over time" (Reynolds, Mavrogenes, Bezruczko, & Hafemann, 1996, p. 1121). The effects of family educational support on children's educational outcomes are usually understood in terms of the theory of parental educational involvement (Epstein, 1988; Grolnick & Slowiaczek, 1994; Patterson, 1986). The essence is that parental involvement makes a child realize the importance of education, which leads to more responsible efforts in school.

After 40 years of struggle with the position that "schools make no difference" (Coleman et al., 1966; Jencks et al., 1972), the research community clearly claims that "schools matter … schools do have major effects upon children's development and to put it simply, schools do make a difference" (Reynolds & Creemers, 1990, p. 1). The literature addresses school effects in terms of school context and school climate. School context describes the nature of student and teacher bodies, as well as the status of material resources of a school. After a systematic review of contextual issues in school effectiveness research, Teddlie, Stringfield, and Reynolds (2000) concluded that school context variables have demonstrated an effect on student educational outcomes.

> The context variable that has been studied the most is the socioeconomic status of the student bodies that attend schools … and the SES makeup of a school has a substantial effect upon student outcomes beyond the effects associated with students' individual ability and social class. (p. 184)

In addition, "the context variables of community type, grade phase, and governance structure have also been demonstrated to have an effect" (p. 184). They finally noted that the effect of school context is more pronounced in the elementary and junior levels than in the senior level. The effects of school context on student educational outcomes are usually understood in terms of the advantages of schools with favorable context such as greater support from parents, fewer disciplinary problems, more positive atmospheres conducive to teaching and learning, and more attraction to highly qualified teachers.

School climate depicts the inner working of a school such as how students are organized for instruction, the expectation that teachers hold for their students, principal's leadership styles, decision-making processes, teachers' classroom practices, and how the school is operated. Citing extensive reviews of school effectiveness literature (Good & Brophy, 1986; Levine & Lezotte, 1995; Purkey & Smith, 1983), D'Agostino (2000) identified orderliness and stability; teacher support; and problem-solving, development, and planning as key indicators of school climate. These indicators speak to disciplinary climate (e.g., student conduct and engagement), principal's leadership (e.g., school missions and protection against outside intrusions), academic pressure (e.g., expectations for students), school culture (e.g., staff turnover and cooperation), teachers' effort (e.g., innovative instructional strategies and community outreach), school policy (e.g., teacher participation and parental involvement in decision making), and continuing education (e.g., professional development for teachers). The effects of school climate on student educational outcomes are usually understood in terms of organizational commitment and continuity (see D'Agostino, 2000, for a review). Essentially, effective schools promote personnel commitment to school missions and motivate staff to achieve the goals of the school. Effective schools also create a formalized structure for stability and predictability and instrumental elements to increase staff responsibility and productivity.

Wang, Haertel, and Walberg (1993, 1994) suggest that the influences of family, school, and society factors on educational outcomes may be determined based on how closely a variable touches on the lives of teachers and students. For example, broad state and district policies are considered the most distal variables, whereas time on task, discipline, and other classroom variables are considered the most proximal variables. Their systematic review identified 228 variables that influence learning that are classified into six categories. The order of influence of these six main variable categories is program design (e.g., curriculum and instruction), out-of-school contextual variables (e.g., home environment, out-of-school activities), classroom instruction and climate (e.g., classroom management), student variables (e.g., motivation, placement), school variables (e.g., parent involvement), and state and district variables (e.g., state education policy). Although it is difficult statistically to quantify the influences of those six categories, Wang et al. (1994) emphasized that the strongest influences come from variables identifiable with classroom management, cognitive and metacognitive processes, student and teacher interactions, and home environmental support. Although there is no direct correspondence with these variables, many composite scales (index variables) developed from the PISA questionnaires can be fairly closely identified with categories of their synthesis.

Our approach in this study involved an application of the input-process-output model of school effectiveness. The concern was how to operationalize this model using questionnaire responses as input and process variables and reading achievement as the output variable. The operationalization was based on the work of Teddlie and Reynolds (2000) and Wang et al. (1993). A close examination of the 2000 PISA questionnaires shows that they measure four principal components (in some cases data aggregation from students to schools

is used): family background, home environment, school context and climate, and school policy and operation. Following the conceptual scheme presented above, data on family background and home environment can be selected to represent inputs, data on school context and climate can be selected to represent educational processes, and data on school policy and operation can be selected to emphasize the policy dimension in educational processes.

Method

Data Source

PISA employed a two-stage stratified random sampling procedure to select a sample from the population of 15-year-old students in each participating country. At the first stage, schools were selected systematically from a comprehensive national list of all eligible schools having 15-year-old students with probabilities proportional to estimated numbers of 15-year-old students enrolled. At the second stage, students were selected in each sampled school from a list of all 15-year-old students. When a school had more than 35 students, 35 students were randomly selected. When a school had fewer than 35 students, all students were selected. As a result, the PISA data had sampling weights for students and schools (Statistics Canada, 2001) that were used at the student and school levels in the present analysis. Adams and Wu (2002) provided details on sample design in the PISA 2000 technical report. The Canadian PISA sample included 29,687 students (at age 15) from 1,117 schools across the 10 provinces (Statistics Canada). Each provincial sample was representative of the population of 15-year-old students in that province.

Variables and Measures

A standardized reading achievement test was administered to students in PISA, in the paper-and-pencil format including multiple-choice, short-answer, and extended-response items (270 minutes of testing time for 141 items). PISA defined reading literacy as "the ability to understand, use, and reflect on written texts in order to participate effectively in life" (Organisation for Economic Cooperation and Development [OECD], 2001, p. 21). This ability was measured through three reading literacy tasks. The first task was retrieving information (locating one or more pieces of information in a text). For example, students were asked to locate information embedded in a play script, a scientific magazine, or a numerical context (e.g., a tree diagram, a graph, or a table). The second task was interpreting texts (constructing meaning by drawing inferences from one or more parts of a text). For example, students were asked to follow the thread of a discussion across paragraphs in a dense text containing strong competing information, to infer an analogical relationship between two phenomena in the text, or to articulate the relationship of information presented in a tree diagram. The last task was reflection and evaluation (relating a text to one's ideas, knowledge, and experience). For example, students were asked to hypothesize about the reason for an author's decision by drawing on evidence in a graph, to evaluate the appropriateness of the ending of a narrative by commenting on its connection with the general theme or model of the text, or to compare claims made in two short texts with their own views and attitudes. OECD (2002) contains a large number of sample questions from the PISA 2000 assessment of reading literacy.

The dependent variable was student reading achievement. This variable took the form of scaled scores (with a mean of 500 and a standard deviation of 100 based on all participating countries) that have been adjusted for reliability, difficulty, and guessing using item response theory (IRT, Statistics Canada, 2001). PISA assigned students to a number of small tests rather than to one large test in order to save time that they would spend on tests. Student reading achievement distribution was then statistically estimated, from which five values (often referred to as plausible values) were randomly drawn for each student. Plausible values for a student are not achievement scores, and they need to be aggregated into a single score for each student. Adams and Wu (2002) provided details on the creation and use of PISA plausible values. Current software programs on multilevel data analysis, Hierarchical Linear Modeling (HLM) in the current case (Raudenbush, Bryk, Cheong, & Congdon, 2000), are able to perform this task. As required statistically, these plausible values were combined in this study as the dependent variable for data analysis.

The independent variables included student-level and school-level characteristics derived from student and school questionnaires and selected on the basis of the conceptual framework discussed above. A total of 25 student-level variables were selected from the student questionnaire, classified into seven categories (see Appendix A), and a total of 22 school-level variables were selected from the school questionnaire, classified into three categories (see Appendix B). Many student-level and school-level variables were index variables. The PISA staff constructed these indices using a number of questionnaire items. Adams and Wu (2002) provided details on the construction of index variables. We rescaled these index variables such that a higher value indicated a more positive response. For the purpose of data analysis, student-level and school-level variables were either standardized to have a mean of zero and a standard deviation of one or were centered around their means. The entire Canadian sample was the reference group for standardizing or centering data.

Statistical Procedures

Multilevel analysis was chosen as the primary statistical technique for our data analysis for two reasons. The first and most obvious was that the data at hand were multilevel in nature (students nested in schools nested in provinces). This data hierarchy must be taken into account in any statistical analysis (Raudenbush & Bryk, 2002). Second, PISA used plausible values to present student academic achievement. Most common statistical packages cannot easily handle plausible values, and statistical packages that can handle plausible values do not allow advanced statistical analysis. Multilevel analysis programs, however, have the option of using plausible values as outcome (dependent) variables.

The ideal multilevel models for this study would include three levels, with students nested in schools nested in provinces. However, the current multilevel programs do not allow weights at the first and second levels simultaneously in a three-level model. This created a technical difficulty because student and school weights were required in analyzing PISA data. To cope with this difficulty, we adopted the strategy of splitting the three-level model into two two-level models. One model had students nested in schools with weights used at the student level. The student-level model was fully developed in this model to produce a posterior mean of reading achievement for each

school with adjustment over student characteristics (student-level variables) in the school. These posterior means of reading achievement, one for each school, were then combined with school-level variables. Based on these school-level variables, the other model had schools nested in provinces, with weights used at the school level.

Although these two two-level models were similar to the three-level model in analytical functions, this strategy achieved simultaneous uses of weights at student and school levels at some costs. One cost was that the partition of variance in reading achievement into components attributable to students, schools, and provinces could not be performed due to the split of the three-level model. The other cost was that the proportion of variance explained at the student, school, and province levels could not be calculated for the same reason. Despite these drawbacks, we did not turn to other statistical methods (e.g., structural equation modeling) because they cannot take data hierarchy into consideration.

These multilevel models helped us identify student and school variables that could account for differences in reading achievement among provinces. To identify provincial characteristics associated with school variables that affected reading achievement, we adopted the technique discussed in Kreft and De Leeuw (1998). The strategy is to set school variables free to vary at the province level and examine the significance of their variation. If these school variables vary significantly at the province level, they are considered as having significant effects on the provincial differences in reading achievement. We adopted this analytic strategy because there were only 10 units at the province level. This small number of units does not permit sophisticated data analysis at the province level (Snijders & Bosker, 1999).

Results

Effects of Student Characteristics

As discussed above, the first two-level model had students nested in schools, and student-level variables were used at the first level of the model to predict reading achievement and function to adjust a school (posterior) mean by student characteristics in that school. When a large number of variables are involved in regression analysis, multicollinearity may occur. We examined correlation among student-level variables to assess multicollinearity. No student-level variables shared correlation large enough to warrant caution. This two-level model was fully developed at the student level, keeping the set of student-level variables significantly (and collectively) related to reading achievement. These results are presented in Table 1.

Among the seven family background variables, six were statistically significant. Because one standard deviation of the PISA reading achievement is 100, it is effortless to convert the effect associated with an independent variable into effect size, standard deviation (*SD*) unit in this case. Rosenthal and Rosnow (1984) classified effect sizes of more than 0.50 *SD* as large, between 0.30 and 0.50 *SD* as moderate, and less than 0.30 *SD* as small. The magnitude of the effects in our study should be understood in the light of these standards. Conventionally, a regression coefficient (effect) is interpreted as the expected change in the dependent variable associated with one unit increase in an independent variable, holding all other independent variables in the model

Table 1
Effects of Student Characteristics on Reading Achievement

Variable	Effect	SE
Family background variables		
Gender (female vs. male)	6.64***	(0.91)
Socioeconomic status (SES)	5.79***	(0.44)
Family structure (single-parent vs. both-parent)	−3.79***	(0.95)
Family size (number of siblings)	−1.82***	(0.34)
Immigration status (immigrant vs. non-immigrant)	−8.40***	(1.65)
Home language (other than English/French vs. English/French)	−10.80***	(1.87)
Home environment variables		
Home educational resources	2.05***	(0.47)
Cultural communication	3.38***	(0.42)
Activities related to classical culture	1.75***	(0.49)
Parental involvement variable		
Family educational support	−7.93***	(0.51)
Reading behavior variables		
Enjoyment of reading	13.44***	(0.66)
Diversity of reading	1.07*	(0.46)
Time spent on reading (half hour vs. zero hour)	2.01**	(0.75)
Time spent on reading (two hours vs. zero hour)	−4.27**	(1.43)
Time spent on reading (three hours vs. zero hour)	−9.23***	(2.42)
Career aspiration variables		
Self-expectation of socioeconomic status	3.74***	(0.46)
Self-expectation of highest education (trade school vs. high school)	9.62***	(2.68)
Self-expectation of highest education (college vs. high school)	19.17***	(2.08)
Self-expectation of highest education (university vs. high school)	25.67***	(2.10)
Academic background variables		
Time spent on homework (in all subjects)	2.30***	(0.42)
Comfort and ability with computers	3.59***	(0.44)
Part-time work variables		
Summer work hours squared	−0.28***	(0.05)
Weekday work hours	−3.39***	(0.95)
Weekday work hours squared	0.35*	(0.17)
Weekend work hours	5.23***	(1.17)
Weekend work hours squared	−1.99***	(0.46)

Note. *$p<0.05$; **$p<0.01$; ***$p<0.001$. N=29,687 students from 1,117 schools. Work hours is used as the number of units with 10 hours as one unit. Among family background variables, family wealth is not statistically significant. Among home environment variables, social communication and possessions related to classical culture are not statistically significant. Among reading behavior variables, time spent on reading (1 hour vs. 0 hour) is not statistically significant. Among academic background variables, school mobility and time spent on homework in language arts are not statistically significant.

constant. Each interpretation should be understood in this manner. If an independent variable is standardized, one unit increase corresponds to one standard deviation increase. This is the case for all index variables in Table 1 (e.g., *higher* and *lower* SES were separated by one standard deviation in SES). We did not make our interpretation this explicit to save space.

Girls scored 7 points higher than boys in reading achievement (0.07 *SD*). Students with higher SES scored 6 points higher than students with lower SES. Students from both-parent households scored 4 points higher than students from single-parent households. A larger number of siblings in a family was related to lower reading achievement, with each additional sibling associated with a decrease in score of 2 points. Non-immigrant students scored 8 points higher than immigrant students. Students with home language as either English or French scored 11 points higher than students with home language other than English and French. Overall, the six family background variables demonstrated small effects on reading achievement.

We found that students from both-parent families scored higher than students from single-parent families. It would be important to examine the SES distribution in these two groups. We did not take this issue into account because we were concerned that such additional modeling might complicate our already complex statistical model (students nested in schools nested in provinces). However, we believe that simpler, specific statistical models should be built in separate investigations (even with the "fishing-trip" approach) to look into this issue.

Among the five variables describing home environment, three were statistically significant. Students with better home educational resources scored 2 points higher in reading achievement than students with poorer such resources. Students with more home cultural communication scored 3 points higher than students with less such communication. Students with more home activities related to classical culture scored 2 points higher than students with fewer such activities. Overall, the three home environment variables demonstrated trivial effects on reading achievement.

Parental involvement is a multidimensional structure of factors (Ho & Willms, 1996). PISA measured only one factor, family educational support, that was statistically significant with a negative effect on reading achievement. Students with less family educational support scored 8 points higher than students with more such support. This finding may indicate two possibilities. First, students with poorer reading achievement received more family educational support than students with better reading achievement. Second, if to some extent family educational support can be considered a form of family academic pressure, then students with more family educational support (as academic pressure) performed worse than students with less such support (or pressure). The negative effect of family educational support is one of the findings in our analysis that has rarely been observed in earlier research. As such it calls for further studies. Although the PISA data do not facilitate a full investigation into our suspicions, a correlation between family educational support and other family characteristics measured in PISA (e.g., home educational resources) may help formulate more specific hypotheses on this negative relationship.

Students with more enjoyment of reading scored 13 points higher in reading achievement than students with less enjoyment of reading. Students who read more diversely scored 1 point higher than students who read less diversely. Reading behavior variables also included four dummy variables denoting time spent on reading daily. It would be desirable to have time spent on reading daily as a continuous measure of time. Unfortunately, PISA constructed this variable as a categorical measure involving zero hour, half hour, one hour, two hours, and three hours. Given the uneven space between neighboring categories, it is impossible to convert this categorical variable into a continuous variable. This is why dummy coding of this categorical variable was used in our analysis. The same situation also existed for self-expectation of highest education for which a similar treatment was taken.

Zero hour reading was the reference category against which half hour, one hour, two hours, and three hours reading were compared. Results indicated the negative effect of overreading. That is, students who spent half an hour reading daily scored 2 points higher than students who spent no time reading; students who spent one hour reading daily scored the same as students who spent no time reading; students who spent two hours reading daily scored 4 points lower than students who spent no time reading; and students who spent three hours reading daily scored 9 points lower than students who spent no time reading.

This clear negative pattern has rarely been observed in earlier research. We suspect that this has to do with the content of reading. Some students read for pleasure (or entertainment); others read for learning (or academic purposes). Reading activities for different purposes may benefit reading achievement differentially. We also suspect that this negative pattern has to do with the speed of reading. Slower readers spend more time on reading (a phenomenon that could be associated with lower reading achievement). Again, the PISA data do not support a full investigation into our suspicions. We recommend that researchers explore the relationships among all reading activity variables, however. For example, correlation between the level of enjoyment of reading and the amount of time spent on reading may inform our suspicions.

Career aspiration variables included three dummy variables measuring self-expectation of highest education. Given that no students expected not to finish high school, high school as the highest education was used as the reference category against which trade school, college, and university as the highest education were compared. Students who expected to complete trade school scored 10 points higher in reading achievement than students who expected to complete high school; students who expected to complete college scored 19 points higher than students who expected to complete high school, and students who expected to complete university scored 26 points higher than students who expected to complete high school. In addition, students who expected higher SES scored 4 points higher than students who expected lower SES. Overall, among all student-level variables, those measuring career aspiration demonstrated the strongest effects on reading achievement.

Two of the four variables describing learning condition were statistically significant. Students who spent more time on homework scored 2 points higher in reading achievement than students who spent less time on home-

work. Students who felt more comfortable with computers scored 4 points higher than students who felt less comfortable with computers. As can be seen, the two learning condition variables demonstrated trivial effects on reading achievement.

Variables descriptive of part-time work were used with their squared terms to examine the accelerated effect of part-time work on reading achievement. Students who worked in the summer scored lower than students who did not work in the summer. For example, students who worked 40, 80, and 120 hours in the summer would have reading achievement 5 points, 19 points, and 41 points lower than students who did not work in the summer. Students who worked on weekdays during the academic year achieved lower than students who did not work on weekdays. For example, students who worked 8, 16, and 32 hours during weekdays would have reading achievement 3 points, 5 points, and 8 points lower than students who did not work during weekdays. Interestingly, students who worked on weekends during the academic year achieved higher than students who did not work on weekends. For example, students who worked 8 hours during weekends would have reading achievement 3 points higher than students who did not work during weekends. Overall, students' part-time working experiences demonstrated fairly important effects on their reading achievement.

In his review of studies on the relationship between academic performance and the number of hours students work per week, Warren (2002) found that results have been consistent: "The number of hours that employed students work per week is negatively related to academic achievement" (p. 366). Results of our analysis in general supported this conclusion. Time spent on summer employment and weekday employment during academic year was associated negatively with reading achievement. However, we did observe an unusual phenomenon about weekend employment. Although the effect on reading achievement in favor of students who worked during weekends (during the academic year) was quite small, it indicated a change of pattern in that weekend employment did not necessarily disadvantage students in reading achievement.

School Effects Across Provinces

As discussed above, the second two-level model had schools nested in provinces, and school-level variables were used at the first level of the model to predict school posterior means in reading achievement estimated from the first two-level model. Table 2 presents the average effects of school-level variables on reading achievement among provinces (termed as average provincial effects on reading achievement). These average provincial effects were estimated with adjustment for student-level variables as reported above, and thus they were over and above individual effects (individual differences) at the student level.

Among the seven school context variables, the percentage of female students in school showed a statistically significant average provincial effect. Across provinces, with a difference of 10% in female enrollment between two schools, the school with the larger percentage of girls achieved 3 points higher in reading achievement than the school with the smaller percentage of girls. No school resources variables had statistically significant average provincial effects. Among the 12 school climate variables, three showed statistically sig-

Table 2
Provincial Effects Associated with School Characteristics
on Reading Achievement

Variable	Average provincial effect	SE	Variation of effect across provinces	Chi-square
School context variables				
School location (urban vs. rural schools)			6.82*	(17.66)
Percentage of female students	2.86**	(1.01)		
Teacher-student ratio			1.50***	(41.24)
School resources variables				
School material resources			4.18**	(23.31)
School instructional resources			3.27*	(18.88)
School climate variables				
Disciplinary climate	4.46**	(1.31)	3.33***	(27.11)
Academic pressure			3.36**	(20.47)
Student behavior			3.67*	(15.40)
Sense of belonging to school	2.69*	(1.08)	2.58*	(18.36)
Teacher morale			2.44*	(15.27)
School autonomy	1.91*	(0.78)		
Instructional minutes in language arts			0.27**	(21.39)
Comparison with national performance			7.37**	(27.69)

Note. *$p<0.05$; **$p<0.01$; ***$p<0.001$. N=1,117 schools from 10 provinces. The relationship between average provincial effect and variation of effect across provinces can be analogized as that between mean and standard deviation. Percentage of female students is used as the number of units with 10% as one unit. Among school context variables, school enrollment size, the percentage of teachers in language arts with a university degree, percentage of teachers participating in professional development programs, and shortage of teachers have neither statistically significant average provincial effects nor statistically significant variation of effects across provinces. Among school resources variables, percentage of computers available to students at age 15 has neither statistically significant average provincial effects nor statistically significant variation of effects across provinces. Among school climate variables, teacher support, teacher behavior, student-teacher relationship, and teacher participation to decision-making have neither statistically significant average provincial effects nor statistically significant variation of effects across provinces. These statistically insignificant estimates are omitted from the table.

nificant average provincial effects. Across provinces, schools with better discipline scored 5 points higher than schools with poorer discipline. Similarly, schools where sense of belonging was stronger scored 3 points higher than schools where sense of belonging was weaker, and schools with greater autonomy scored 2 points higher than schools with weaker autonomy.

Although these average provincial effects appear small, they are practically important in that, as mentioned above, these effects were over and above individual differences in reading achievement. They point to directions of important improvement that will benefit all provinces in their efforts to improve the reading literacy of their students and thus provide a general

guideline for school improvement that leads to improvement in reading achievement in each province.

Differential Provincial Effects

Each province has an effect associated with a school-level variable. Examining whether the average provincial effect of this school-level variable is statistically significant among provinces is only one part of the analysis. Table 2 also indicates whether the effects of school-level variables on reading achievement varied across provinces (termed as differential provincial effects on reading achievement). School-level variables whose effects vary statistically significantly across provinces are considered provincial characteristics because variation in provincial effects across provinces is more indicative of whether a school characteristic is a provincially critical variable (Kreft & De Leeuw, 1998). We examined, therefore, not only average provincial effects, but also variation in provincial effects across provinces.

Two of the seven school context variables were statistically significantly different in their effects on reading achievement across provinces. School location is a categorical variable in PISA that describes the population of the area where a school is located. These categories include village (population fewer than 3,000), small town (3,000-15,000), town (15,000-100,000), city (100,000-1,000,000), and big city (more than 1,000,000). In our analysis, the first three categories were combined to denote rural, and the last two were combined to denote urban. The effects of both school location and teacher-student ratio on reading achievement varied significantly across provinces. This indicates that school location and teacher-student ratio were significant provincial variables. Among the three school resources variables, the effects of both school material resources and school instructional resources on reading achievement varied statistically significantly across provinces, indicating that school material resources and school instructional resources were also significant provincial variables.

Seven of the 12 school climate variables showed statistically significant variation in their effects on reading achievement across provinces. The effects of disciplinary climate; academic pressure; student behavior; sense of belonging to school; teacher morale; instructional minutes in language arts; and use of comparison with district, provincial, and national performance on reading achievement varied significantly across provinces. These variables were deemed significant provincial variables on reading achievement.

Such a large number of significant provincial (school-related) variables indicates that provinces are differentially successful in using school effects to promote provincial reading achievement of their students. This finding provides opportunities for provincial intervention to improve student reading achievement as a whole. It also calls for exchange of experiences among provinces to learn from one another's success and failure. The Council of Ministers of Education Canada (2000) is in a good position to promotion discussions among provinces.

Table 3 further illustrates the differential provincial effects of school-level variables on reading achievement across provinces by calculating and tabulating provincial (posterior) effects for provinces (from the second two-level model with schools nested in provinces). An estimate in Table 3 indicates the

Table 3

Provincial Effects of School Characteristics on Reading Achievement

	NL	PEI	NS	NB	QC	ON	MB	SK	AB	BC
School context variables										
School location (urban vs. rural schools)	2.77	0.80	0.02	0.92	–1.91	3.75	–7.45	1.66	–10.87	15.11
Teacher-student ratio	0.15	–0.31	0.35	–1.14	–1.25	3.39	–0.35	0.02	0.11	0.10
School resources variables										
School material resources	1.43	0.07	0.85	–1.08	–4.76	0.09	1.59	–0.13	0.29	1.01
School instructional resources	–0.12	–1.07	1.38	–2.31	–7.12	–3.35	1.37	0.94	–1.56	3.33
School climate variables										
Disciplinary climate	3.75	6.37	5.81	4.19	7.14	12.22	4.18	4.19	4.36	12.08
Academic pressure	–3.87	–2.16	–2.10	1.22	–5.41	2.14	–4.27	–1.76	–1.52	–8.29
Student behavior	1.21	1.34	0.32	–4.09	2.76	12.20	5.15	7.18	5.68	5.79
Student sense of belonging to school	2.63	2.94	0.91	0.35	6.22	5.81	3.55	4.76	3.68	4.16
Teacher morale	0.84	0.65	1.55	3.02	3.29	11.09	–0.27	0.31	1.05	–2.38
Instructional minutes in language arts	3.21	1.46	2.22	–1.23	0.57	–0.49	–0.09	–0.87	3.19	0.74
Comparison with national performance	2.83	1.62	4.30	–7.12	16.95	1.31	0.44	–1.87	–10.68	–5.41

Note. School characteristics in the table have statistically significant variation of effects on reading achievement across provinces. Instructional minutes in language arts is used as the number of units with 10 minutes as one unit.

relationship between reading achievement and a given school-level variable in a province. This table does not provide a way to compare provinces in terms of higher or lower reading achievement, but rather a measure of how this school-level variable is differentially related to reading achievement across provinces. Therefore, Table 3 measures the variability of this relationship (or effect) across provinces.

Specifically, for each school-level variable, a provincial (posterior) mean effect on reading achievement was calculated to represent the provincial effect of this school-level variable in each province. Some school-level variables varied in their effects on reading achievement across provinces. These variations were tested statistically. Differential provincial effects were offered in a descriptive sense; that is, we could not perform any post-hoc analysis on provincial effects. It is likely that only two provinces have significantly different provincial effects among the 10 provinces. Therefore, the interpretation of Table 3 was entirely descriptive even though comparisons were made among provinces.

Table 3 shows that in seven provinces (Newfoundland, Prince Edward Island, Nova Scotia, New Brunswick, Ontario, Saskatchewan, and British Columbia), urban schools performed better than rural schools. In three provinces (Quebec, Manitoba, and Alberta), rural schools performed better than urban schools. In six provinces (Newfoundland, Nova Scotia, Ontario, Sas-

katchewan, Alberta, and British Columbia), schools with a lower teacher-student ratio (a teacher teaches fewer students) performed better than schools with a higher teacher-student ratio (a teacher teaches more students). In four provinces (Prince Edward Island, New Brunswick, Quebec, and Manitoba), schools with a higher teacher-student ratio performed better than schools with a lower teacher-student ratio.

School material resources had positive effects on reading achievement in seven provinces (Newfoundland, Prince Edward Island, Nova Scotia, Ontario, Manitoba, Alberta, and British Columbia). Schools with more material resources performed better than schools with fewer material resources. Three provinces (New Brunswick, Quebec, and Saskatchewan) had negative effects. In terms of the effect of school instructional resources on reading achievement, four provinces (Nova Scotia, Manitoba, Saskatchewan, and British Columbia) had positive effects. Schools with more instructional resources performed better than schools with fewer instructional resources. Six provinces (Newfoundland and Labrador, Prince Edward Island, New Brunswick, Quebec, Ontario, and Alberta) had negative effects.

In all provinces, disciplinary climate had positive effects on reading achievement. Schools with better discipline performed better than schools with poorer discipline. Academic pressure had positive effects on reading achievement in two provinces (New Brunswick and Ontario). Schools with higher academic pressure performed better than schools with lower pressure. Eight provinces (Newfoundland, Prince Edward Island, Nova Scotia, Quebec, Manitoba, Saskatchewan, Alberta, and British Columbia) had negative effects.

In terms of the effect of student behavior on reading achievement, nine provinces (except for New Brunswick) had positive effects. Schools with better student behavior performed better than schools with poorer student behavior. Only New Brunswick had a negative effect. Sense of belonging to school had positive effects on reading achievement in all provinces. Schools where sense of belonging was stronger performed better than schools where sense of belonging was weaker.

In terms of the effect of teacher morale on reading achievement, eight provinces (except for Manitoba and British Columbia) had positive effects. Schools with better teacher morale performed better than schools with poorer teacher morale. Manitoba and British Columbia had negative effects. The effect of instructional minutes in language arts on reading achievement was positive in six provinces (Newfoundland, Prince Edward Island, Nova Scotia, Quebec, Alberta, and British Columbia). Schools with more instructional time in language arts performed better than schools with less instructional time in language arts. Four provinces (New Brunswick, Ontario, Manitoba, and Saskatchewan) had negative effects.

In terms of the effect on reading achievement associated with principals' making comparisons with district, provincial, and national performance (to help develop school policies and practices), six provinces (Newfoundland, Prince Edward Island, Nova Scotia, Quebec, Ontario, and Manitoba) had positive effects. Schools where administrators compared school performance with district, provincial, and national performance achieved higher than schools

where administrators did not make such comparison. Four provinces (New Brunswick, Saskatchewan, Alberta, and British Columbia) had negative effects.

For a number of school-level variables, we have counter-intuitive findings in some provinces. Instead of considering them as abnormal results, we would emphasize the uniqueness of these provinces. Something made teachers or students react to certain questionnaire items differently in these provinces than in others. For example, New Brunswick is the only province that showed a negative effect of student behavior on reading achievement. Teachers in high performing schools in that province might take a unique (e.g., tougher) standard on student behavior, counting behavior tolerable in other schools as misbehavior. In such a case, the negative effect actually reflects less tolerance of student misconduct in those schools. In another example, we found better performance of rural than urban schools in three provinces. The dividing line between urban and rural is the population of 100,000. Many schools that we labeled rural might be suburban schools in those provinces. In general, the literature has shown that students in suburban schools tend to outperform students in urban (inner city) schools.

Finally, although many provincial effects in Table 3 were small, the focus of that table was on the variability, not the magnitude of provincial effects. For example, teacher-student ratio had small effects in most provinces. Still, Ontario showed a relatively large positive effect (3.39) and Quebec a relatively large negative one (–1.25). It is this variability that is captured in Table 3.

Discussion

The common way to discuss provincial differences in reading achievement is to introduce province-level variables to explain variance in reading achievement among provinces. This approach was considered unfruitful for this study because of the small number of units (cases) at the province level (10 provinces). We expect that with such low statistical power, it would not be possible to detect any significant province-level effects. Therefore, instead of using province-level variables to explain provincial differences in reading achievement, we concentrated on schools in each province. The idea was to examine how schools in various provinces produced differential effects (through school resources, context, and climate) on reading achievement. By examining how differently, for example, school climate was related to reading achievement across provinces, we attempted to separate provinces where differences in school climate accounted for differences in reading achievement.

For example, as shown in Table 3, although disciplinary climate had a positive effect on reading achievement in every province, provinces were significantly different in their schools' ability to relate disciplinary climate to reading achievement. Therefore, provincial differences in reading achievement were related to their schools' disciplinary climate. The implication is that province-wide improvement in school disciplinary climate is likely to be associated with improvement in student reading achievement as a whole in a province. Therefore, not only is our strategy methodologically sounder given our data circumstances, but also this strategy is able to generate unique implications for educational policies and practices. We provide several policy implications in this study. Some are closely aligned with suppositions and speculations that could go beyond the data. We use these assertions to provoke

discussion and debate on education policy issues that may affect students' academic achievement. In so doing, we might invite caution, alternative ideas, or further review from research perspectives different from ours. We also caution that our analytical results do not infer any causal relationship. Although we adopted words such as *effects* and *impacts* in our interpretation and discussion as commonly used when reporting regression analysis, regression analysis (multilevel regression analysis in our case) indicates association rather than causation.

We emphasize two school-level variables: disciplinary climate and sense of belonging to school. They showed significant average province effects on reading achievement, and their effects varied significantly across provinces. It is our belief that if provinces wish to improve their students' reading achievement (perhaps academic achievement in general), disciplinary climate and students' sense of belonging should be on their working agenda with their schools. If student behavior (another important school-level variable shown in our study) can be considered part of school discipline, there are even more reasons to be concerned about disciplinary climate.

The results for disciplinary climate are consistent with those from other studies (DeBaryshe, Patterson, & Capaldi, 1993). There are obvious advantages to improving disciplinary climate in schools. Although disciplinary climate is largely a school matter, school policies in this area need to be supported by provincial policies and legislations that permit principals and teachers to engender a positive disciplinary environment. For example, the Edmonton City Council has passed a local bylaw that prohibits bullying in school. This means to students that bullying incidents in Edmonton schools are now subject to the investigation of law reinforcement agencies. Provinces that take on this type of legislation to ensure a positive disciplinary climate in their schools are in a promising position to improve the academic achievement of their students.

As for student sense of belonging to school, our results suggest that school policies that encourage identification with the school may relate positively to reading achievement. Quebec may be a good place to illustrate the importance of sense of belonging. For whatever political motives, Quebec seemed to be successful in creating student sense of belonging that showed the largest positive effect on reading achievement across all provinces. We speculate that perhaps it is the nationalism that rallies Quebec students around a common sense of belonging. If there is any educational lesson in this finding, it is that one way to promote student sense of belonging is for each province to have a political, cultural, or social theme or pride that can be commonly shared among educators, parents, and students, and the resulting sense of belonging among students may be effectively associated with their academic achievement as shown in this study.

Quebec stands out in the effects of school material and instructional resources, with both variables showing a relatively large negative effect. One possible clue to this phenomenon may lie in the 1999 SAIP Science results (Council of Ministers of Education Canada, 2000), where Quebec schools reported fewer constraints on instruction due to limitations on instructional resources than schools in other jurisdictions. We suspect that the (relative) lack of school material and instructional resources may have become an advantage rather

than a disadvantage for Quebec teachers in that it stimulates them to seek alternative options to facilitate their instruction. We suggest that it is teachers rather than materials that matter for differences in academic achievement. Either as a suspicion or common wisdom, we believe that a working hypothesis can be formulated and open for investigation that such a policy orientation may eventually benefit student academic achievement. We are not calling for a funding reduction in school material and instructional resources. We are suggesting that provinces may give priority to teachers rather than resources if they cannot accommodate both teachers and resources in their educational systems (in times of economic difficulties, for example).

The effect of academic pressure (or pressure to achieve) on reading achievement was negative in most provinces. This result is not consistent with other studies that suggest that academic emphasis is positively associated with academic achievement (Ma & Klinger, 2000). But it does raise a concern about whether academic pressure is unreasonably high in some Canadian schools. We think that another explanation may be more appropriate, however, in the light of the superior performance of Canadian students in reading internationally. That is, the negative effects of academic pressure in those provinces may indicate that teachers were paying close attention to students whose performance was satisfactory from an international perspective, but still below the local academic expectation. No matter whether our speculation is true or not, we think that in general attention to students at the lower academic achievement distribution regardless of their international standings should be either rewarded and encouraged to continue if it is present in Canadian schools or promoted if not present. It is common wisdom that such attention can only improve the academic performance of Canadian students as a whole.

The differential effects for Alberta and New Brunswick on one hand and Quebec on the other have significant implications in the light of the strong emphasis on provincial assessments in all these provinces and the increasing interest everywhere in interprovincial and international comparisons. Our speculation is that educators in Quebec may be less skeptical about the value of such information than those in other jurisdictions. If so, then there are important lessons to be learned from how Quebec uses evaluation information. Even if it is not so, we believe that a general suggestion can be warranted that comparative information can be used positively, for example, to celebrate the achievement of high performing schools (rather than to highlight the problems of low performing schools) or to show those outside a given province what has been accomplished. We hypothesize that such attitudes toward comparative evaluation boost teacher and student morale to achieve an even higher standard of performance.

We believe that the advantage associated with the percentage of female students in a school represents an aggregated phenomenon of female advantage in reading achievement at the student level. The implication is not to create single-sex schools, but that provinces must seriously address this gender inequity in language arts that exists in most provinces (Human Resources Development Canada et al., 2001). This situation reminds us of the several decades of struggle or effort in eliminating gender differences in mathematics achievement that were once so prevalent. Mathematics educators have started

to reap the harvest of this effort, with several meta-analytical studies showing that gender differences in mathematics achievement have been narrowing (Friedman, 1996; Frost, Hyde, & Fennema, 1994). We hypothesize that the same can be achieved with gender differences in reading achievement.

There is no shortage of debate on which is the better educational policy: central control or school autonomy. In the current age of emphasis on central control (e.g., consolidating local school boards, changing the funding balance in the direction of more provincial and less local funding), our results regarding school autonomy are provocative. We found a significant average provincial effect that a high degree of school autonomy was associated with a superior level of reading achievement. This demonstrates that school autonomy does play an undeniable role if our goal is to improve the academic achievement of students as a whole. We hypothesize that provinces that seriously balance between school autonomy and central control may be in a good position to advance the academic achievement of their students.

In sum, we detected quite a few critical school-level variables whose effects on reading achievement varied significantly across provinces. We found that most of these school-level variables described school climate, together with a couple of school-level variables descriptive of school context and school resources. This finding is encouraging in that it is school climate rather than school context or school resources that educators can reform meaningfully (or that is under the direct control of educators). In this sense, our study does provide Canadian educators and policymakers with the motivation to effect educational change.

We realize that we are dealing with a complex set of data with students nested in schools nested in provinces, and we made compromises in our statistical approach. We suggest that these data be reexamined when multilevel programs advance further. We also realize that analyses of other national databases may lead to different outcomes, and in due course a comprehensive synthesis of existing databases may prove to be necessary. As we mention, some policy implications in this study depart somewhat from our analytical results, and so they are tentative but provocative assertions or hypotheses. We invite researchers to join a discussion that will lead to further investigations that generate finer evidence-based working knowledge for Canadian educators and policymakers.

References

Adams, R., & Wu, M. (Eds.). (2002). *Programme for International Student Assessment (PISA): PISA 2000 technical report*. Paris: Organisation for Economic Cooperation and Development.

Arzubiaga, R., Rueda, R., & Monzo, L. (2002). *Family matters related to the reading engagement of Latina/o children*. Ann Arbor, MI: University of Michigan, Center for the Improvement of Early Reading Achievement.

Coleman, J.C. (1990) *Foundations of social theory*. Cambridge, MA: Harvard University Press.

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

Council of Ministers of Education Canada. (2000). *Science learning: The Canadian context*. Toronto, ON: Author.

Crocker, R.K. (2002a). A decade of SAIP: What have we learned? What do we need to know? In P. DeBroucker & A. Sweetman (Eds.), *Towards evidence-based policy for Canadian education* (pp. 207-225). Ottawa: Statistics Canada.

Crocker, R.K. (2002b). *Learning outcomes: A critical review of the state of the field in Canada*. Ottawa: Canadian Educational Statistics Council.

Croll, P. (2004). Families, social capital and educational outcomes. *British Journal of Educational Studies, 52*, 390-416.

D'Agostino, J.V. (2000). Instructional and school effects on students' longitudinal reading and mathematics achievements. *School Effectiveness and School Improvement, 2*, 197-235.

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.

Entwisle, D.R., & Alexander, K.L. (1988). Factors affecting achievement test scores and marks of Black and White first graders. *Elementary School Journal, 88*, 449-472.

Epstein, J.L. (1988). How do we improve programs for parent involvement? *Educational Horizons, 66*, 75-77.

Friedman, L. (1996). Meta-analysis and quantitative gender differences: Reconciliation. *Focus on Learning Problems in Mathematics, 18*, 123-128.

Frost, L.A., Hyde, J.S., & Fennema, E. (1994). Gender, mathematics performance, and mathematics-related attitudes and affect: A meta-analytic synthesis. *International Journal of Educational Research, 21*, 373-386.

Gallimore, R., & Goldenberg, C. (1993). Activity settings of early literacy: Home and school features in children's emergent literacy. In E.A. Forman, N. Minick, & C.A. Stone (Eds.), *Contexts for learning: Sociocultural dynamics in children's development* (pp. 315-335). New York: Oxford University Press.

Good, T.L., & Brophy, J.E. (1986). School effects. In M.C. Wittrock (Ed.), *Handbook of research in teaching* (pp. 570-602). New York: Macmillan.

Grolnick, W.S., & Slowiaczek, M.L. (1994). Parents' involvement in children's schooling: A multidimensional conceptualization and motivational model. *Child Development, 65*, 237-252.

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

Horvat, E., Weininger, E., & Lareau, A. (2003). From social ties to social capital. *American Educational Research Journal, 40*, 319-352.

Human Resources Development Canada, Statistics Canada, & Council of Ministers of Education Canada. (2001). *Measuring up: The performance of Canada's youth in reading, mathematics and science*. Ottawa: Authors.

Jencks, C., & Phillips, M. (1998). The Black-White test score gap: An introduction. In C. Jencks & M. Phillips (Eds.), *The Black-White test score gap* (pp. 1-52). Washington, DC: Brookings Institution.

Jencks, C.S., Smith, M., Acland, H., Bane, M.J., Cohen, D., Ginitis, H., Heyns, B., & Michelson, S. (1972). *Inequality: A reassessment of the effect of family and schooling in America*. New York: Basic Books.

Kreft, I., & De Leeuw, J. (1998). *Introducing multilevel modeling*. Thousand Oaks, CA: Sage.

Levine, D.U., & Lezotte, L.W. (1995). Effective schools research. In J.A. Banks & C.A. McGee Banks (Eds.), *Handbook of research on multicultural education* (pp. 525-547). New York: Macmillan.

Ma, X., & Klinger, D.A. (2000). Hierarchical linear modelling of student and school effects on academic achievement. *Canadian Journal of Education, 25*, 41-55.

Organisation for Economic Cooperation and Development. (2001). *Knowledge and skills for life: First results from PISA 2000*. Paris: Author.

Organisation for Economic Cooperation and Development. (2002). *Programme for international student assessment: Sample tasks from the PISA 2000 assessment of reading, mathematical and scientific literacy.* Paris: Author.

Patterson, G.R. (1986). Performance models for antisocial boys. *American Psychologist, 41*, 432-444.

Podell, D.M., & Soodak, L.C. (1993). Teacher efficacy and bias in special education referrals. *Journal of Educational Research, 86*, 247-253.

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

Purkey, S.C., & Smith, M.S. (1983). Effective schools: A review. *Elementary School Journal, 83*, 427-452.

Raudenbush, S.W., & Bryk, A.S. (2002). *Hierarchical linear models* (2nd ed.). Thousand Oaks, CA: Sage.

Raudenbush, S.W., Bryk, A., Cheong, Y.F., & Congdon, R. (2000). *Hierarchical linear and nonlinear modeling*. Lincolnwood, IL: Scientific Software International.

Reynolds, D., & Creemers, B. (1990). School effectiveness and school improvement: A mission statement. *School Effectiveness and School Improvement, 1*, 1-3.

Reynolds, A.J., Mavrogenes, N.A., Bezruczko, N., & Hafemann, M. (1996). Cognitive and family-support mediators of preschool effectiveness: A confirmatory analysis. *Child Development, 67*, 1119-1140.

Robitaille, D.F., & Taylor, A.R. (2000). *TIMSS-Canada report, Volume 5: New findings for a new century*. Vancouver, BC: University of British Columbia Press.

Robitaille, D.F., Taylor, A. R., & Orpwood, G. (1996). *TIMSS-Canada report, Volume 1: Grade 8*. Vancouver, BC: University of British Columbia Press.

Robitaille, D.F., Taylor, A.R., & Orpwood, G. (1997). *TIMSS-Canada report, Volume 2: Grade 4*. Vancouver, BC: University of British Columbia Press.

Rogoff, B. (1990). *Apprenticeship in thinking: Cognitive development in social context*. Oxford, UK: Oxford University Press.

Rosenthal, R., & Rosnow, R.L. (1984). *Essentials of behavioral research: Methods and data analysis*. New York: McGraw-Hill.

Schneider, B., & Stevenson, D. (1999). *The ambitious generation: America's teenagers, motivated but directionless*. New Haven, CT: Yale University Press.

Snijders, T.A.B., & Bosker, R.J. (1999). *Multilevel analysis*. Thousand Oaks, CA: Sage.

Statistics Canada. (2001). *OECD PISA user's manual*. Ottawa: Author.

Snow, C.E., Barnes, W.E., Chandler, J., Goodman, I.F., & Hemphill, L. (1991). *Unfulfilled expectations: Home and school influences on literacy*. Cambridge, MA: Harvard University Press.

Teddlie, C., & Reynolds, D. (Eds.). (2000). *The international handbook of school effectiveness research*. London: Falmer.

Teddlie, C., Stringfield, S., & Reynolds, D. (2000). Context issues within school effectiveness research. In C. Teddlie & D. Reynolds (Eds.), *The international handbook of school effectiveness research* (pp. 160-186). London: Falmer.

Tharp, R.G., & Gallimore, R. (1988). *Rousing minds to life: Teaching, learning, and schooling in social context*. New York: Cambridge University Press.

Wang, M.C., Haertel, G.D., & Walberg, H.J. (1993). Toward a knowledge base for school learning. *Review of Educational Research, 63*, 249-294.

Wang, M.C., Haertel, G.D., & Walberg, H.J. (1994). What helps students learn? *Educational Leadership, 51*(4), 74-79.

Warren, J.R. (2002). Reconsidering the relationship between student employment and academic outcomes: A new theory and better data. *Youth and Society, 33*, 366-393.

Washington, J.A. (2001). Early literacy skills in African-American children: Research consideration. *Learning Disabilities Research and Practice, 16*, 213-221.

Weisner, T.S. (1984). Ecocultural niches of middle childhood: A cross cultural perspective. In W.A. Collins (Ed.), *Development during childhood: The years from six to twelve* (pp. 335-369). Washington, DC: National Academy of Science Press.

Appendix A
Variables at the Student Level

Family Background Variables

Gender (female = 1 vs. male = 0), socioeconomic status (economic, social, and cultural status), family structure (single-parent = 1 vs. both-parent = 0), family size (number of siblings), immigration status (immigrant = 1 vs. non-immigrant = 0), home language (language other than English/French = 1 vs. English/French = 0), family wealth (availability of a dishwasher, a room of their own, educational software, internet; number of cellular phones, TV sets, computers, motor cars, bathrooms).

Home Environment Variables

Home educational resources (availability and number of dictionaries, desks and quiet places for study, textbooks, calculators), cultural communication (discussing political or social issues, books, films or TV programs, listening to classical music), social communication (discussing, eating together, spending time talking), home possessions related to classical culture (availability of

classical literature, books of poetry, works of art), home activities related to classical culture (visiting museums or art galleries; attending operas, ballets or classical symphony concerts; watching live theatre).

Parental Involvement Variable
Family educational support (doing school work with fathers, mothers, brothers, sisters).

Reading Behavior Variables
Enjoyment of reading (not giving reading up, reading in spare time, getting totally absorbed when reading), diversity of reading (magazines, comic books, fiction, non-fiction, email and Web pages, newspapers), time spent on reading (half hour = 1 vs. zero hour = 0), (one hour = 1 vs. zero hour = 0), (two hours = 1 vs. zero hour = 0), (three hours = 1 vs. zero hour = 0).

Career Aspiration Variables
Self-expectation of socioeconomic status, self-expectation of highest education (trade school = 1 vs. high school = 0), (college = 1 vs. high school = 0), (university = 1 vs. high school = 0).

Learning Condition Variables
School mobility (attending different schools), time spent on homework, and comfort and ability with computers.

Part-Time Work Variables
Summer work hours, weekday work hours, and weekend work hours.
Note. Coding information for dummy variables is provided. All other variables are continuous indices.

<center>

Appendix B
Variables at the School Level

</center>

School Context
School enrollment size, school location (city schools = 1 vs. country schools = 0), percentage of female students, teacher-student ratio, percentage of teachers in language arts with a university degree, percentage of teachers participating in professional development programs, and shortage of teachers.

School Resources
School material resources (condition of buildings as well as heating, cooling, lighting systems; instructional space), school instructional resources (computers, instructional materials in the library, multi-media resources, science laboratory equipment, facilities for the fine arts), percentage of computers available to students at age of 15.

School Climate
Disciplinary climate (teachers waiting long for students to quiet down, students not working well, students not listening to teachers, students not starting working long after lessons begin, noise and disorder in classroom), academic press (teachers wanting students to work hard, telling students that they can do better, not liking it when students deliver careless work; students having to learn a lot), teacher support (teacher helping students with their work, continuing teaching until students understand, doing a lot to help students, helping

students with their learning), student behavior (absenteeism, disrupting classes, skipping classes, lacking respect for teachers, using alcohol or illegal drugs, intimidating or bullying others), teacher behavior (absenteeism, low expectations, poor student-teacher relations, not meeting individual students' needs, resisting change, too strict with students, students not being encouraged to achieve their full potential), student-teacher relationship (students getting along well with teachers, teachers being interested in students' well-being, teachers listening to what students have to say, students getting extra help, teachers treating students fairly), sense of belonging to school (feeling like an outsider, making friends easily, feeling like belonging, feeling awkward and out of place, feeling being liked, feeling lonely), teacher morale (high morale, working with enthusiasm, taking pride in their schools, valuing academic achievement), teacher participation to decision making (number of categories as not being teacher responsibilities), school autonomy (number of categories as not being school responsibilities), instructional minutes in language arts, and comparison with district, provincial, and national performance (yes = 1 vs. no = 0).

Note. Coding information for dummy variables is provided. All other variables are continuous indices.