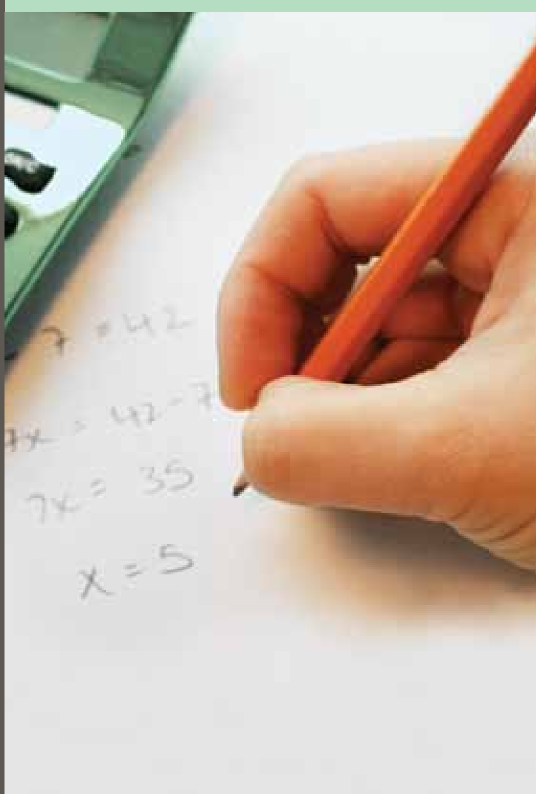
Pan-Canadian Assessment Program

PCAP-2010

Contextual Report on Student Achievement in Mathematics



cmec
Council of Ministers of Education, Canada
Conseil des ministres de l'Éducation (Canada)

Pan-Canadian Assessment Program

PCAP-2010

Contextual Report on Student Achievement in Mathematics

cmec

Council of Ministers of Education, Canada

Conseil des ministres de l'Éducation (Canada)

The Council of Ministers of Education, Canada (CMEC) was formed in 1967 by the jurisdictional ministers responsible for education to provide a forum in which they could discuss matters of mutual interest, undertake educational initiatives cooperatively, and represent the interests of the provinces and territories with national educational organizations, the federal government, foreign governments, and international organizations. CMEC is the national voice for education in Canada and, through CMEC, the provinces and territories work collectively on common objectives in a broad range of activities at the elementary, secondary, and postsecondary levels.

Through the CMEC Secretariat, the Council serves as the organization in which ministries and departments of education undertake cooperatively the activities, projects, and initiatives of particular interest to all jurisdictions. One of the activities on which they cooperate is the development and implementation of pan-Canadian testing based on contemporary research and best practices in the assessment of student achievement in core subjects.

Note of appreciation

The Council of Ministers of Education (Canada) would like to thank the students, teachers, and administrators whose participation in the Pan-Canadian Assessment Program ensured its success. The quality of your commitment has made this study possible. We are truly grateful for your contribution to a pan-Canadian understanding of educational policy and practices in mathematics, science, and reading among Grade 8 students.

Council of Ministers of Education, Canada
95 St. Clair West, Suite 1106
Toronto, Ontario M4V 1N6

Telephone: (416) 962-8100
Fax: (416) 962-2800
E-mail: cmec@cmec.ca

ISBN 978-0-88987-219-6

Ce rapport est également disponible en français.

Printed on recycled paper.

TABLE OF CONTENTS

LIST OF TABLES AND CHARTS

1 What Is the Pan-Canadian Assessment Program?

2 Overview of Achievement Results

3 Student, Teacher, and School Characteristics

4 Student Attitudes

6 Instructional Climate

7 Time Allocation and Use

8 Teaching and Learning Strategies in Mathematics

9 Assessment

10 Population Differences and Achievement Equity

11 Summary Model, Robust Effects, and Further Research

EXECUTIVE SUMMARY

The Pan-Canadian Assessment Program (PCAP) 2010 is the continuation of CMEC's commitment to inform Canadians about how well their education systems are meeting the needs of students and society. The information gained from this pan-Canadian assessment provides ministers of education with a basis for examining the curriculum and other aspects of their school systems. PCAP is administered every three years to a sample of more than 30,000 Grade 8 students (Secondary Two in Quebec),[1] representative of provinces and territories and of the two official language groups within these jurisdictions.

Three subjects — reading, mathematics, and science — are assessed in each cycle, with one subject being treated as a major domain and the other two as minor domains. The major domain is assessed in greater detail than the minor domains. Reading was the major domain in 2007 and mathematics the major domain in 2010. In addition to the student tests, questionnaires are administered to students, teachers, and school principals. These questionnaires are designed to measure demographic and socioeconomic factors and to gather information about attitudes, school policies and practices, and teaching and learning strategies.

The results of each assessment are published in two major reports. The first is a public report, which gives scores on the major and minor subjects by jurisdiction, language, and gender. The second is a contextual report, which examines achievement in the major domain in relation to variables derived from the questionnaires. This executive summary gives highlights of the 2010 Contextual Report focusing on mathematics.

The first two chapters of this report present an introduction to PCAP and a brief summary of the mathematics achievement results. Chapters 3 through 9 give results for clusters of questionnaire variables and their relationship to mathematics achievement. Three main sets of results are given. First, responses to specific questionnaire items are summarized by population.[2] Next, simple relationships between specific variables and mathematics scores are presented. These are referred to as simple regression relationships because they link two variables, the response to a single questionnaire variable and average mathematics scores. Finally, at the end of each chapter, a multiple regression model is presented, which examines the relationship of each of the questionnaire variables examined in the chapter, while controlling for all of the other variables in that particular cluster.

[1] The first PCAP cycle in 2007 used a sample of 13-year-olds. This was changed to a sample of Grade 8 students in 2010 and in subsequent assessments to simplify administration of the assessments.

[2] The term "population" refers to the combination of jurisdiction (province/territory) and official language group within each jurisdiction because these are the subgroups from which the samples were drawn.

The following table gives highlights of the variables examined in each of Chapters 2 through 9. The most robust results are those given in Chapters 10 and 11. A more detailed summary of these two chapters is given below.

Chapter	Topic	Variables Examined
2	Overview of achievement results	Differences in mathematics scores by jurisdiction and language
3	Student, teacher, and school characteristics	Demographic and socioeconomic characteristics of students; school demographics; teacher qualifications and experience
4	Student attitudes	Attitudes toward school; attitudes toward mathematics; confidence with mathematics; attributions of success and failure
5	Student mathematics behaviours and strategies	Strategies for dealing with difficult mathematics problems; out-of-school activities; early mathematics learning
6	Instructional climate	Areas of emphasis in mathematics; class size; influences on school programs; accommodation of special-needs students; challenges in teaching mathematics
7	Time allocation and use	Time on mathematics; student absence; homework time; types of homework
8	Teaching and learning strategies in mathematics	Instructional strategies; resources
9	Assessment	Classroom assessments; external assessments; uses of assessments; attitudes toward assessment

Rather than attempting to summarize all of the individual results, the focus here is on the final two chapters. Chapter 10 is concerned with achievement differences among populations and how these differences relate to equity. Equity was examined in two ways. First, the amount of variation in scores was compared for each population. This showed that few populations had a score pattern that was much more variable than the Canadian average. However, several had lower variability, which can be taken as having achieved higher equity in scores. For example, the variability for British Columbia French, Saskatchewan English, Prince Edward Island,[3] and Alberta French was about 20 per cent or more lower than that for Canada as a whole. A plot of equity versus achievement for populations revealed that greater equity is associated with higher average population scores. Some populations thus come closer than others to meeting the goal of both high achievement and high equity.

In a second stage of analysis, each population was compared to a reference population (Ontario English in most cases) in a model designed to determine if the populations differ in the way their scores are influenced by the questionnaire variables. This analysis shows that the average mathematics achievement of most populations, relative to Ontario English, improves once other variables are controlled. Based on this model, it can be argued that the comparatively high score for Ontario English is thus a function of some characteristics of the Ontario population, of which the most obvious examples are student demographics, school demographics, and student attitudes and attributions.

Chapter 11 presents a composite model, designed to identify robust results — those that are stable across the simple and multiple regression models. Taken together, all of the variables used in the final model account for 41 per cent of the variability in student scores and 73 per cent of the variability in school average scores. Statistically significant simple and multiple regression effects were found for a large number of variables. No single variable stands out as having a decisive influence. However, several clusters of variables show larger effects than others. Student demographic characteristics and attitudes have larger and more consistent effects on student average achievement than do teaching and learning strategies. Student attitudes, along with school demographic characteristics also have the largest effects on school average achievement. Most of the other variable clusters have statistically significant but smaller effects.

[3] Because of the small number of francophone students in the sample for Prince Edward Island, the two language groups are not separated.

The table below highlights what we have called the robust effects. In statistical terms, these are the unique or residual effects of specific variables after controlling for all of the other variables in the model. These variables can thus be said to influence mathematics achievement independently of any of the other variables in the model.

Variables associated with higher achievement in mathematics	Variables associated with lower achievement in mathematics
Student level variables	
• Higher expected education level • More books in the home • Higher mother's education • Use of informal early mathematics learning activities • Likes school • Attitude that mathematics is easy • Attribution of success and failure to ability • General confidence in mathematics • Confidence with computers and calculators • Persistence in dealing with difficult mathematics problems • Personal communications • Direct instruction • Uses calculators in mathematics • Strategic approach to mathematics learning • Total homework time • Absence for school-related reasons • Conventional assessment	• Uses English or mostly English in a variety of contexts inside and outside school • Uses a language other than English or French in a variety of contexts inside and outside school • Negative attitudes to mathematics • Attribution of success or failure to luck (fatalism) • Decreased confidence in mathematics ability over time • Use of on-line help in mathematics • Seeks outside help in mathematics • Uses technology for entertainment • Indirect instruction • Instruction by projects/assignments • Graphic/pictorial learning strategies • Seeks help with mathematics • Mathematics homework time • Absence for non–school-related reasons • Unconventional assessment
School/teacher level variables	
• Larger school size • Private school • Larger class size • More teacher-assigned homework time	• Adult other than a teacher in the classroom • Percentage of students absent for non–school-related reasons • Time lost because of student misbehaviour • Negative view of jurisdictional assessments

The results might be interpreted as showing that the characteristics students bring to schools and the structural features of schools themselves have greater influence on mathematics achievement than what is done in schools or classrooms. However, the results also highlight a design issue with cross-sectional surveys of this nature. In effect, the student and school characteristics measured are relatively stable and permanent, whereas the teaching and learning variables are likely more transient. Most of the teaching and learning variables capture, at best, what might have been done during the school year in which the assessment was conducted. This is not necessarily indicative of the student's exposure to teaching and learning over the years of schooling up to Grade 8. On the other hand, this cumulative exposure is obviously reflected in the PCAP mathematics test.

It is beyond the scope of an omnibus report such as this to investigate all of the possible links among the variables in the model and how these affect mathematics achievement. For example, it was not possible in this report to look at other research related to the variables examined. The design of PCAP provides for a research phase that would follow the release of the public and contextual reports. Research work of this kind can extend the findings of this report in relation to policy issues. These include structural features of the school system such as class size and time allocations, as well as teaching and learning variables such as homework, teaching strategies, and assessment practices. Even a cursory examination of previous large-scale surveys such as PCAP 2007 and the SAIP and PISA studies reveals that there are many consistent patterns among the variables affecting achievement in mathematics and other subjects. Before drawing any strong policy conclusions from many of the effects seen here, an effort should be made to determine if these results are consistent with what has been found in other similar studies. The ability to replicate results is key to scientific research, and having consistent results can greatly strengthen any policy decisions that might be made from these results.

1 WHAT IS THE PAN-CANADIAN ASSESSMENT PROGRAM?

The Pan-Canadian Assessment Program (PCAP) 2010 is the continuation of CMEC's commitment to inform Canadians about how well their education systems are meeting the needs of students and society. The information gained from this pan-Canadian assessment provides ministers of education with a basis for examining the curriculum and other aspects of their school systems.

School curriculum programs vary from jurisdiction to jurisdiction across the country, so comparing results from these varied programs is a complex task. However, young Canadians in different jurisdictions learn many similar skills in mathematics, reading, and science. PCAP has been designed to determine whether students across Canada reach similar levels of performance in these core disciplines at about the same age, and to complement existing assessments in each jurisdiction so they have comparative Canada-wide data on the achievement levels attained by Grade 8 students across the country.

Goals

When the ministers of education began planning the development of PCAP in 2003, they set out the following goals for a conceptually new pan-Canadian instrument of assessment designed to:

- inform educational policies to improve approaches to learning;
- focus on mathematics, reading, and science, with the possibility of including other domains as the need arises;
- reduce the testing burden on schools through a more streamlined administrative process;
- provide useful background information using complementary context questionnaires for students, teachers, and school administrators; and
- enable jurisdictions to use both national and international results to validate the results of their own assessment programs and to improve them.

Table 1-1 provides CMEC's actual and proposed dates for administering PCAP to Canadian Grade 8 students.

Table 1-1 **Actual and prospective PCAP administrations**

Domains	Spring 2007 (13-year-olds)	Spring 2010 (Grade 8 students)	Spring 2013 (Grade 8 students)
Major	Reading	Mathematics	Science
Minor	Mathematics	Science	Reading
Minor	Science	Reading	Mathematics

The development process

In August 2003, a PCAP working group of experienced and knowledgeable representatives from several jurisdictions — including an external authority on measurement theory, large-scale assessment, and educational policy, began the development process. A concept paper was commissioned that would elaborate on issues of structure, development planning, operations, and reporting. Drawing on this concept paper, the working group defined PCAP as a testing program that would:

- be administered at regular intervals to students who are 13-year-olds at the start of the school year;
- be based on the commonality of all current jurisdictional curricular outcomes across Canada;
- assess mathematics, science, and reading;
- provide a major assessment of one domain, with a minor concentration on the two other domains; and
- focus on reading as the major domain in the first administration in 2007, mathematics in 2010, and science in 2013.

As of 2010, it was determined that PCAP would be administered to Grade 8 students, and, whenever possible, intact classes were selected in order to minimize the disruption to classrooms and schools.

For each subject area, a thorough review of curricula, current assessment practices, and research literature was then undertaken, and reports were written to indicate the common expectations among all jurisdictions.

The working groups for bilingual framework development, established for each of the three subject areas, were composed of representatives from several jurisdictions with knowledge and experience in curriculum and assessment for the particular subject. Each working group also had an external expert in the assessment of the particular subject to advise and assist with the development of a framework statement establishing the theory, design, and performance descriptors for each domain. The framework statements were reviewed and accepted by all participating jurisdictions as the basis for test-item development.

Bilingual teams for developing the test items were then established; members of these teams were subject-area educators selected from all jurisdictions, with a subject-assessment expert to supervise. Each subject framework provided a blueprint with its table of specifications describing the subdomains of each subject area, the types and length of texts and questions, the range of difficulty, and the distribution of questions assessing each specific curriculum expectation.

Texts and questions were developed in both official languages and cross-translated to be equivalent in meaning and difficulty. Jurisdictions reviewed and confirmed the validity of the French-English translations to ensure fair and equitable testing in both languages.

All new items were reviewed by outside validators and further revised by members of the item-development team. These texts and items were then submitted to the framework-development working group to be examined in light of the blueprint, and field-test booklets were consequently put together. Booklets contained both selected-response and constructed-response items. Their range of difficulty was deemed accessible to Grade 8 students, based on scenarios meaningful to the age group and reflecting Canadian values, culture, and content.

Field testing involved the administration of these temporary forms to a representative sample of students from an appropriate range of jurisdictions in both languages. Approximately 2,000 students in 100 schools across Canada were involved in the field testing. The tests were then scored by teams of educators from the jurisdictions. Following analysis of the data from the field test, each framework-development working group reviewed all items and selected the texts and items considered best, from a content and statistical viewpoint, to form four 90-minute booklets.

Design and development of contextual questionnaires

The accompanying questionnaires for students, teachers, and schools were designed to provide jurisdictions with contextual information that would contribute to the interpretation of the performance results. Such information could also be examined and used by researchers, policy-makers, and practitioners to help determine what factors influence learning outcomes.

A questionnaire-development group comprising educators and research experts from selected jurisdictions developed a framework to ensure that the questions asked of students, teachers, and school principals were consistent with predetermined theoretical constructs or important research questions. The group:

- reviewed models of questionnaire design found in the three large-scale assessment programs — the School Achievement Indicators Program (SAIP); the International Association for the Evaluation of Educational Achievement (IEA) – Trends in International Mathematics and Science Study (TIMSS); and the Programme for International Student Assessment (PISA);
- maximized research value by shaping the questionnaires around selected research issues for the 2010 administration of the test.

The questionnaires were adapted and expanded for mathematics as the major domain.

This report focuses on the questionnaire results and especially on relationships between questionnaire responses and scores on the mathematics assessment.

Features of the administration of the PCAP 2010 mathematics assessment

In the spring of 2010, the test was administered to a random sample of schools and Grade 8 classes (one per selected school) with a random assignment of booklets.

Sampling

This assessment adopted the following stratified sampling process in the selection of participants:

1. the random selection of schools from each jurisdiction, drawn from a complete list of publicly funded schools provided by the jurisdiction;
2. the random selection of Grade 8 classes, drawn from a list of all eligible Grade 8 classes within each school;
3. the selection of all students enrolled in the selected Grade 8 class;
4. when intact Grade 8 classes could not be selected, a random selection of Grade 8 students.

> The sampling process refers to the way in which students were selected to write the assessment. It is necessary to select a large enough number of participants to allow for adequate representation of the population's performance; the word "population" refers to all eligible students within a jurisdiction and/or a linguistic group.

In the case where numbers were smaller than the desired size, all schools and/or all Grade 8 classes meeting the criteria within the jurisdiction were selected. This approach ensured that we had an adequate number of participants to allow for reporting on their achievement as if all students within the jurisdiction had participated.

The sampling process resulted in a very large sample of approximately 32,000 Grade 8 students participating in the assessment. All students answered questions in all three domains. Approximately 24,000 responded in English, and 8,000 in French.

Reporting results by language

The results obtained from students educated in the French system of their respective jurisdiction are reported as French. The results obtained from students educated in the English system of their respective jurisdiction are reported as English. Results achieved by French immersion students who wrote in French are calculated as part of the English results since these students are considered to be part of the English-language cohort. All French and English students were expected to write for 90 minutes, with breaks deemed appropriate by the assessment administrator. If necessary, students were given an additional 30 minutes to complete the assessment. Then, they completed the context questionnaire at the back of their test booklet.

Participation

Each school received the assessment handbook that outlined the purposes of the assessment, the organization and administration requirements, and suggestions to encourage the maximum possible participation. These suggestions included a common administration script to ensure that all students encountered the testing process in a similar manner, and provided guidelines for accommodating special-needs students. PCAP testing is intended to be as inclusive as possible in order to provide a complete picture of the range of performance for students in Grade 8. The students who were excused from participating were nevertheless recorded for statistical purposes; they included those with functional disabilities, intellectual disabilities, socio-emotional conditions, or limited language proficiency in the target language of the assessment.

Participation rates

In large-scale assessments, participation rates are calculated in a variety of ways and are used to guide school administrators when determining whether the number of students who completed the assessment falls within the established norm set for all schools. In the case of PCAP, a formula for this purpose is provided to the test administrators, thereby ensuring that all schools use the same guidelines and that the set minimum of participating students is uniformly applied. Using this formula, the overall PCAP student participation rate was over 85 per cent. For additional information concerning student participation and sampling, refer to Chapter 2.

Schools were encouraged to prepare and motivate students for the test, aiming for positive participation and engagement in the process by teachers, students, and parents. The materials provided included information pamphlets for parents and students, as well as the school handbook.

Schools were also asked to have the teacher questionnaire completed by all the mathematics teachers of the participating students in the school, and the school questionnaire by the school principal. All questionnaires were linked to student results but used unique identifiers to preserve confidentiality.

Scoring the student response booklets

The scoring was conducted concurrently in both languages in one location over a three-week period. After all student booklets had been submitted from the jurisdictions, the booklets were then scrambled into bundles of 10 so that any single bundle contained booklets from several jurisdictions. The scoring-administration team, the table leaders, and the scorers themselves came from several jurisdictions. The whole scoring process included:

- parallel training of both table leaders and scorers in each subject area;
- a bilingual committee with responsibility for reviewing all instruments and selecting anchor papers to ensure comparability at every level;
- twice-daily rater-reliability checks in which all scorers marked the same student work in order to track the consistency of scoring on an immediate basis; and
- double-blind scoring in which 300 of each of the four booklets were returned to the scoring bundles to be re-scored, providing an overall inter-rater reliability score.

Structure of this report

The PCAP 2010 Public Report, released in November 2011 (CMEC, 2011), presented detailed performance results. Chapter 2 gives a brief summary of these results.

The main focus of this report is on the questionnaire results and specifically on variables associated with mathematics achievement. The report is divided into chapters corresponding to major clusters of variables that, according to previous research and theory, may influence mathematics scores. These include demographic characteristics; attitudes and attributions; early mathematics learning; student out-of-school activities and behaviours; instructional climate; teaching and learning strategies; allocation and use of time; and assessment practices.

In each of Chapters 3 to 9, questionnaire results are first presented descriptively by jurisdiction and language. This is followed by a two-stage analytic process. First, the mathematics scores are compared for students across categories on each of the variables of interest. These comparisons are used to determine whether the variable is significantly associated with mathematics achievement. Second, the relationships between questionnaire variables and achievement are examined through a multilevel regression modelling process designed to allow the effects of a single variable to be examined while controlling for other variables within the same cluster.

Chapter 10 examines the data from the perspective of differences between populations and achievement equity. The main issue in this chapter is whether the data can be used to account for the observed differences in mathematics achievement across jurisdictions and official languages. Also, since the variation in mathematics achievement is greater in some populations than in others, possible reasons for these differences are explored as well.

Finally, Chapter 11 presents a summary model in which each variable showing significant results in earlier models is re-examined with all other variables controlled. This model is intended to identify “robust” effects — those that remain statistically significant even when all other variables in the model are controlled.

2 OVERVIEW OF ACHIEVEMENT RESULTS

Statistical Note

Samples. The results presented in this report are based on samples. Separate samples were selected for each jurisdiction (province or territory) and for anglophone and francophone populations within each jurisdiction. Some of the francophone samples were quite small. Because statistics such as percentages or means are quite unstable for small samples, it was necessary to combine the two language groups in some jurisdictions when reporting results at the jurisdictional level. For student results, the language groups were combined for Prince Edward Island, Newfoundland and Labrador, and Yukon. For teacher and school data, the language groups were also combined for British Columbia, Alberta, and Nova Scotia because the numbers of teachers and schools were much smaller than the numbers of students. Students in French immersion programs were considered part of the anglophone population. When pan-Canadian results were computed, all students, schools, and teachers were assigned to their appropriate language group.

Confidence intervals. The results from the samples are estimates of those that would have been achieved had all members of the populations been included in the assessment. The actual results may differ from their population values for a variety of reasons, including sampling error or unreliability in responses to test or questionnaire items. It is common practice in surveys of this sort to report a range within which the actual population value is expected to fall. This range is known as a **confidence interval**. Confidence intervals are reported in tables as a number with a ± (plus or minus) sign, which represents the range above or below the reported value in which the population value is expected to be found with a specified level of probability, typically 95 per cent. Confidence intervals are represented in bar graphs by **error bars**, which correspond to the 95 per cent confidence interval above and below the number given by the bar. We can say that the population value would be expected to be within the range represented by the total width of the error bars, 95 times out of 100.

Statistical significance. When making comparisons between groups (such as the difference in mean mathematics scores for jurisdictions), the difference is said to be statistically significant if the observed difference is greater than the sum of the two confidence intervals. For graphical presentations, a difference can be considered **statistically significant** if the error bars for the groups being compared do not overlap. To keep the graphs as simple as possible, statistical significance is indicated in this report mainly for comparisons of mean mathematics scores across groups and for regression coefficients. However, confidence intervals for all results are given in the tables presented in the appendix.

Weights. The ratio of population to sample size gives a statistic called the **weight**, which is applied when results are combined across groups. This ensures that each population or sub-population is represented in the combined results in proper proportion to its size. All results given in this report use weighted data, so the results can be said to represent the whole population. However, error computations are based on actual sample sizes, as errors are strongly related to sample size.

Populations and samples

The sampling process was described in Chapter 1. Table 2-1 gives the student, school, and teacher sample sizes for each jurisdiction and the official-language groups within jurisdictions.[4] The small sample sizes for some of the francophone populations led to a decision to combine the language groups in Prince Edward Island, Newfoundland and Labrador, and Yukon.

All students wrote all three domains of the assessment, and all completed the questionnaires, so all student results are based on the complete sample.

TABLE 2-1 **Samples***

	Student Sample	School Sample	Teacher Sample
British Columbia (E)	3,328	147	325
British Columbia (F)	231	11	15
Alberta (E)	3,183	145	147
Alberta (F)	332	22	23
Saskatchewan (E)	2,838	149	155
Saskatchewan (F)	80	7	8
Manitoba (E)	2,788	150	153
Manitoba (F)	322	15	16
Ontario (E)	3,374	144	181
Ontario (F)	2,509	142	148
Quebec (E)	1,703	87	104
Quebec (F)	3,534	130	143
New Brunswick (E)	1,611	89	89
New Brunswick (F)	1,053	62	61
Nova Scotia (E)	2,548	136	139
Nova Scotia (F)	295	10	10
Prince Edward Island	484	25	28
Newfoundland and Labrador	1,861	122	128
Yukon	305	10	17
Canada	32,379	1,603	1,890

* Population numbers are not provided here. Because they are derived from the sum of the weights in the data file which are based on the actual number of students, teachers, and schools completing the questionnaire, these numbers may differ slightly from the PCAP 2010 Technical Report. The teacher sample was based on the school and student samples. All teachers who taught mathematics to students writing the PCAP test in a school were sampled. Because intact classes were used, one teacher was sampled in most schools, with two or more teachers in a few schools. For purposes of standard error computations, in this report, teacher population estimates are based on school weights.

[4] The combination of jurisdiction and language is referred to in this report as a population, since this is the primary unit from which schools and students were sampled. Most of the comparisons given in this report are between populations.

Scaling

Following the initial scoring process, as described in Chapter 1, scores were scaled to a mean of 500 and a standard deviation of 100 for Canada. This provides a relatively simple basis for comparing groups. On this type of scale, approximately two-thirds of the individual student scores will fall within plus or minus one standard deviation of the mean, or between 400 and 600.

Overview of achievement results

Chart 2-1 gives mean mathematics scores for the jurisdictions. This shows that Quebec and Ontario students perform at a level significantly above the Canadian average and those in Alberta are at the Canadian average, while students in all other jurisdictions perform below the Canadian average.

Chart 2-2 shows the results for the two official-language groups for each jurisdiction for which a breakdown is possible. The picture for the anglophone populations is similar to that shown in Chart 2-1. On the francophone side, none of the populations is above the Canadian average; Quebec, Ontario and New Brunswick are at the Canadian average; and all others are below the Canadian average. In this case, the Quebec French population makes up such a large proportion of the Canadian French average that Quebec will almost always be at or near the Canadian French average.

CHART 2-1 **Mean mathematics scores and confidence intervals by jurisdiction**

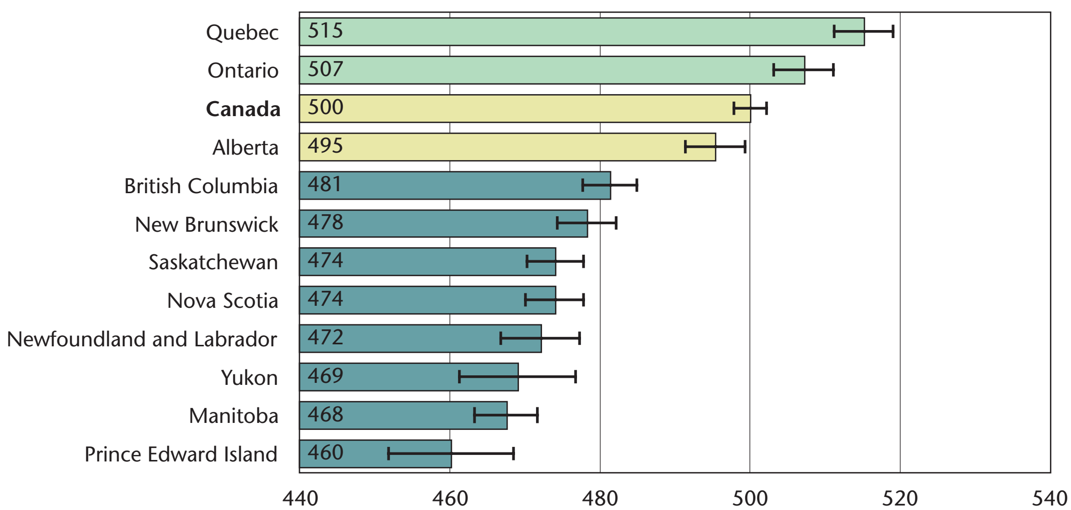
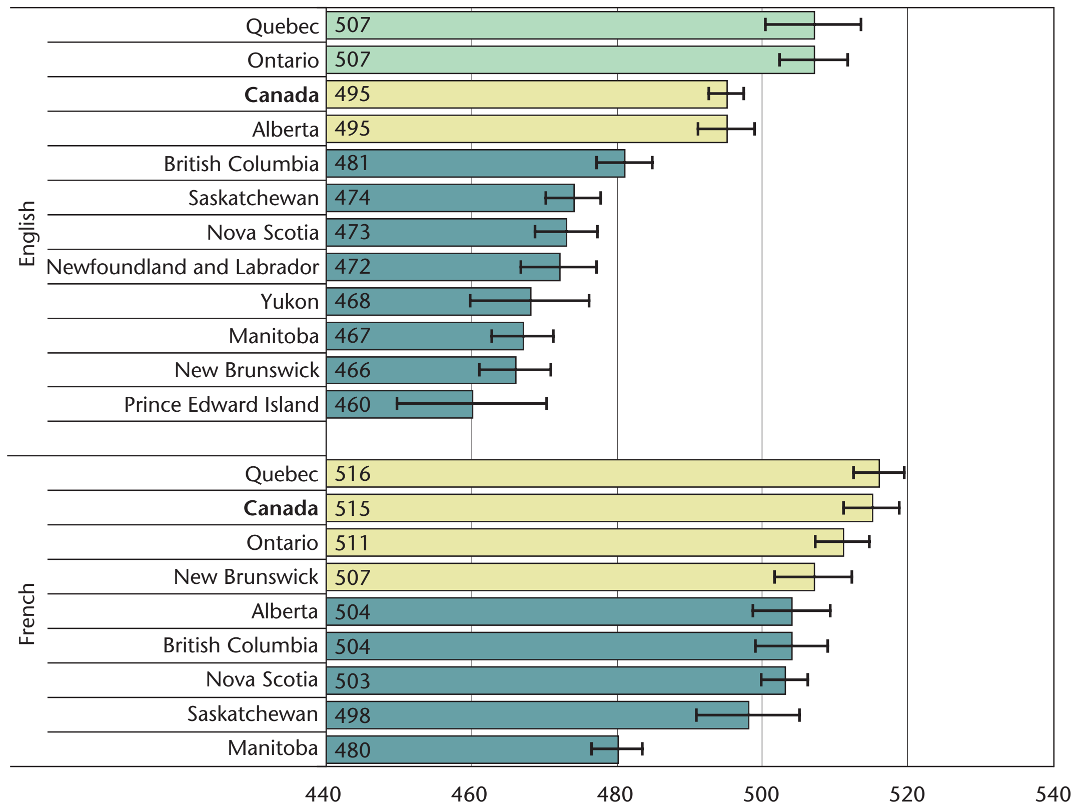


CHART 2-2 **Mean mathematics scores and confidence intervals by jurisdiction and language**

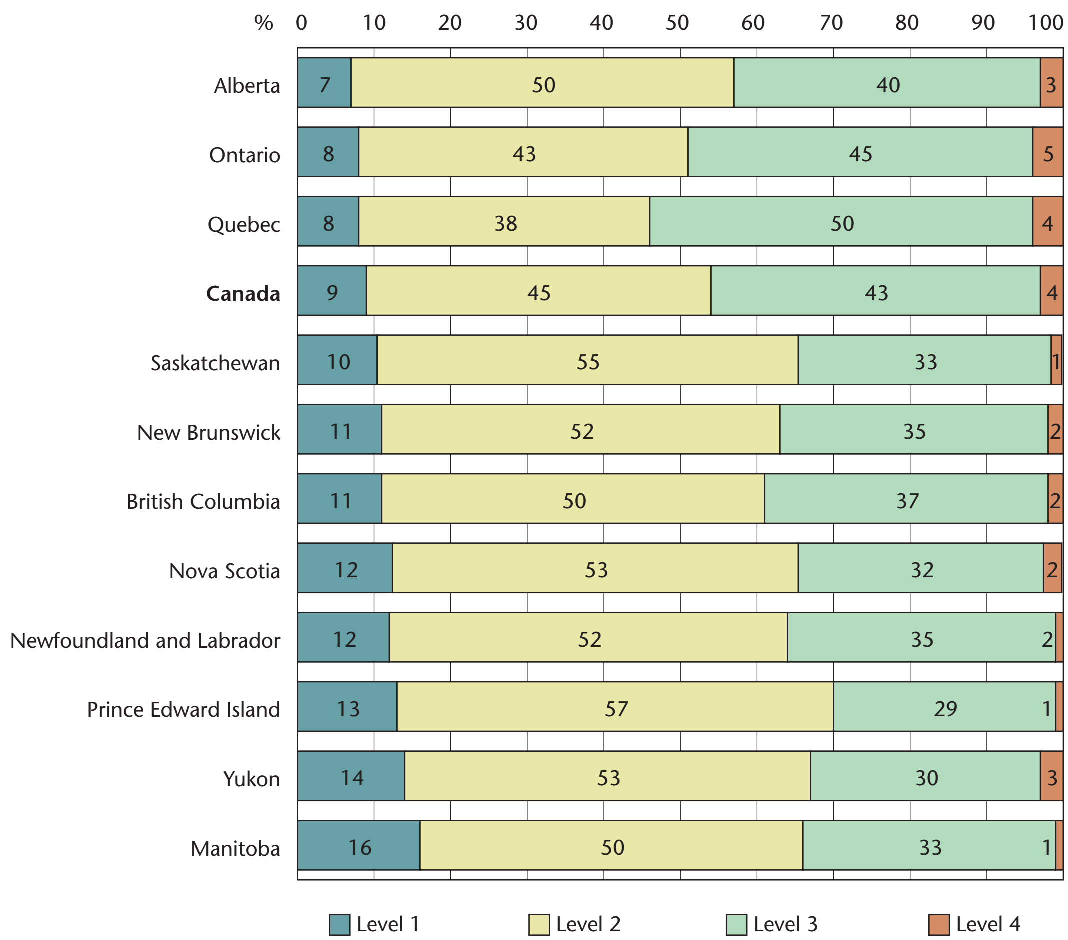


Mathematics proficiency levels

Another way of looking at mathematics performance is to establish proficiency levels based on descriptions of what students can do at each level. For the mathematics test, four proficiency levels were defined, with level 2 considered the acceptable level of performance for Grade 8 students.[5] Performance levels were then summarized as the percentage of students reaching each level.

The results by jurisdiction are given in Chart 2-3. Canada-wide, 91 per cent of students perform at the acceptable level or higher. Fewer than 20 per cent of students perform below the acceptable level in any jurisdiction. However, the range for level 1 performance varies considerably, from 7 per cent in Alberta to 16 per cent in Manitoba. Relatively few students in any jurisdiction are at level 4, the highest level, from 5 per cent in Ontario to 1 per cent in Saskatchewan, Prince Edward Island, and Manitoba.

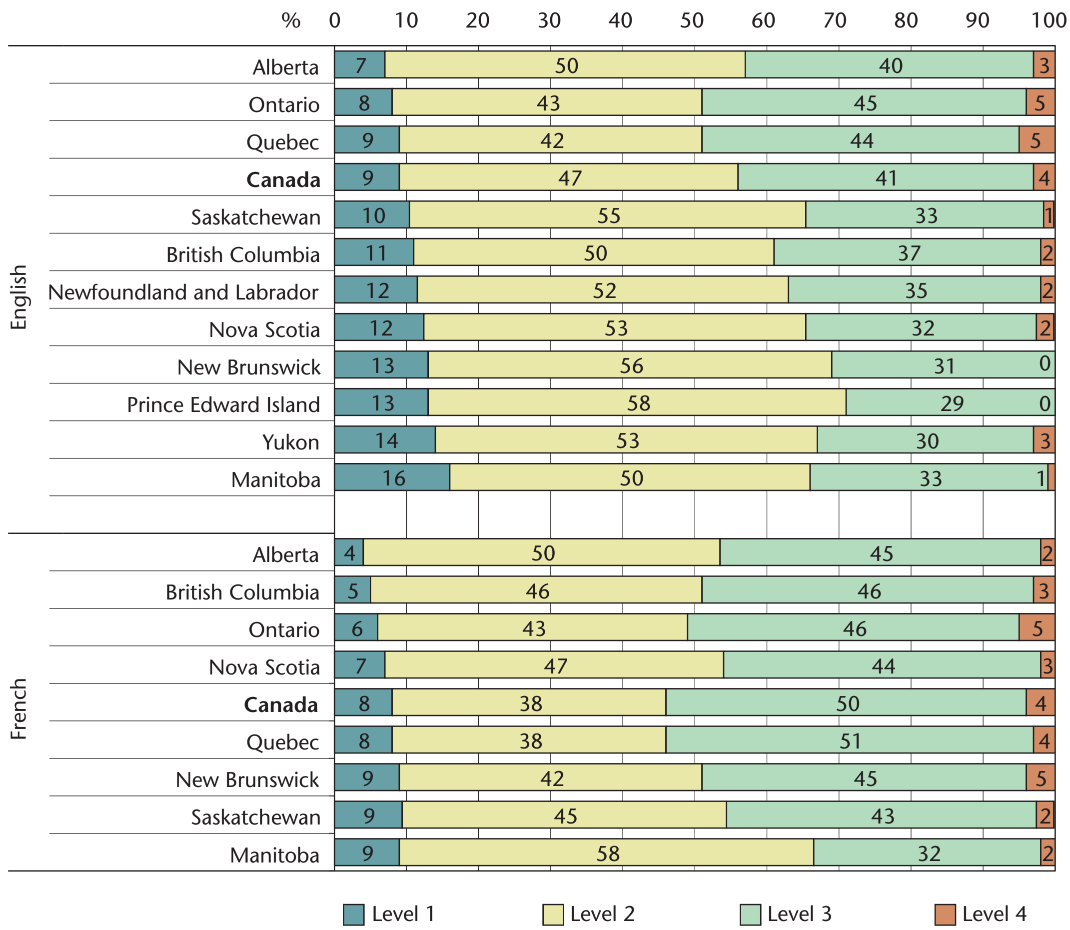
CHART 2-3 **Mathematics proficiency levels by jurisdiction**



[5] For details on the level definitions, please see the PCAP 2010 public report (**http://www.cmec.ca/Publications/Lists/Publications/Attachments/274/pcap2010.pdf**).

Chart 2-4 shows the results for mathematics proficiency levels by jurisdiction and language. The pattern here is similar to that for the other graphs, with somewhat wider variations among the English than among the French populations.

CHART 2-4 **Mathematics proficiency levels by jurisdiction and language**



Summary

In terms of the scale used, differences between jurisdictions amount to approximately 75 points, or three-fourths of a standard deviation on the scale. Differences within language groups are somewhat smaller than this, at 47 points for anglophone populations and 36 points for francophone populations. In terms of proficiency levels, level 2 was considered the minimum for satisfactory performance. Canada-wide, 91 per cent of students reached that level or higher, ranging from 93 per cent in Alberta to 84 per cent in Manitoba. More detailed information about the achievement results are reported in the public report: *PCAP 2010: Report on the Pan-Canadian Assessment of Mathematics, Science, and Reading.*

3 STUDENT, TEACHER, AND SCHOOL CHARACTERISTICS

This chapter presents demographic and socioeconomic characteristics of students, teachers, and schools. These are considered as fixed characteristics of individuals and of the system and are thus treated in the models as antecedent conditions to teaching and learning. To facilitate comparisons, descriptive/comparative results for the selected variables are reported by jurisdiction and language. Analytical results related to mathematics achievement are presented in two forms. First, comparisons are given for mathematics proficiency levels and mean scores across categories for each of the variables of interest. Second, these characteristics are entered into a regression equation as predictors of achievement in mathematics and the results presented in terms of regression coefficients. Regression analysis allows the effect of each variable to be examined while controlling for other variables in the model. More detail on how the regression analysis was conducted is given later in this chapter.

In subsequent parts of this report, the background variables used in this section are treated as covariates[6] as they are generally not within the control of the school system, and their influence on achievement is considered largely independent of educational policy or practice. That is not to say that the system should not take account of these variables. In particular, since an important goal of schooling is to promote equity, it is appropriate to develop policies that can help overcome any disadvantage created by socioeconomic or other background characteristics.

Student characteristics

Student gender

Charts 3-1 and 3-2 give the gender distribution of students by jurisdiction and language. Generally, the proportions of males and females would be expected to depart from the expected 50 per cent each only by random amounts, based on sampling error. However, statistically significant differences were found in several jurisdictions.[7] In particular, the overall proportions of males in minority francophone and anglophone populations and in Yukon were less than expected. The opposite is true for Prince Edward Island, where the proportion of males is higher than in any other population.

[6] A covariate is a variable that is entered into a regression equation to act as a control against other variables of more direct interest. The effects for the variables of interest are computed after controlling for the covariates.

[7] To keep the graphs as simple as possible, error bars are not plotted on 100 per cent graphs in this report. However, tables giving the relevant standard errors may be found in the PCAP 2010 Technical Report. Any references to statistical significance reported here may be confirmed by consulting these tables.

CHART 3-1 **Males and females by jurisdiction**

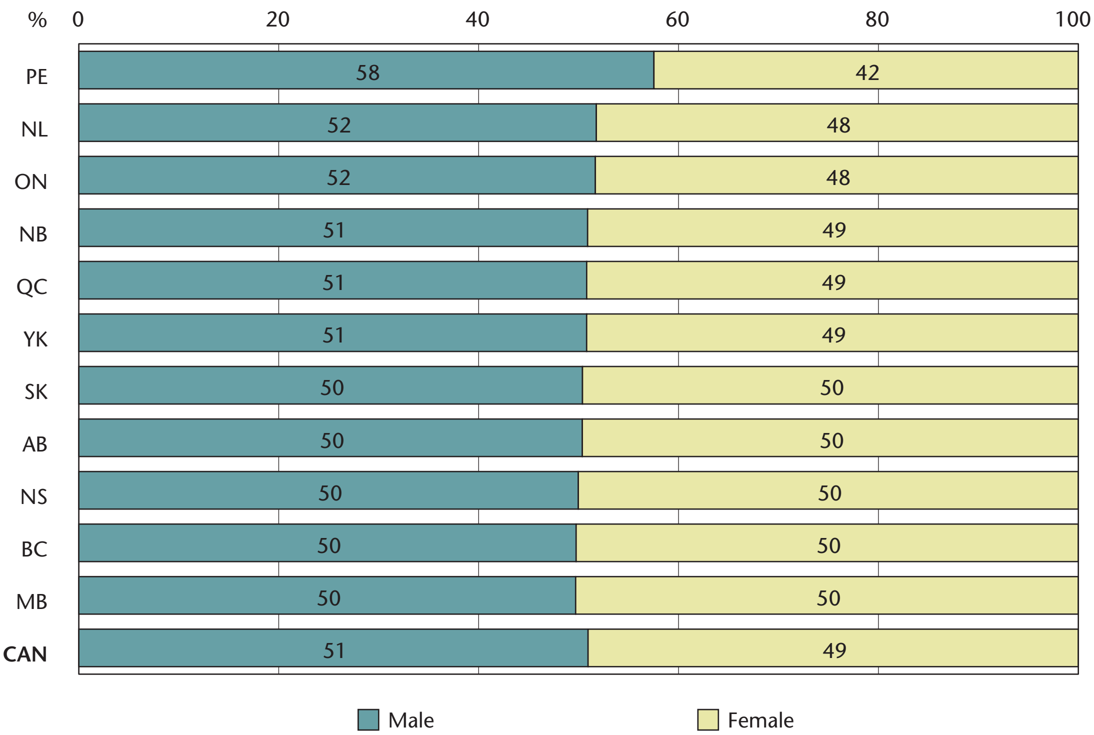
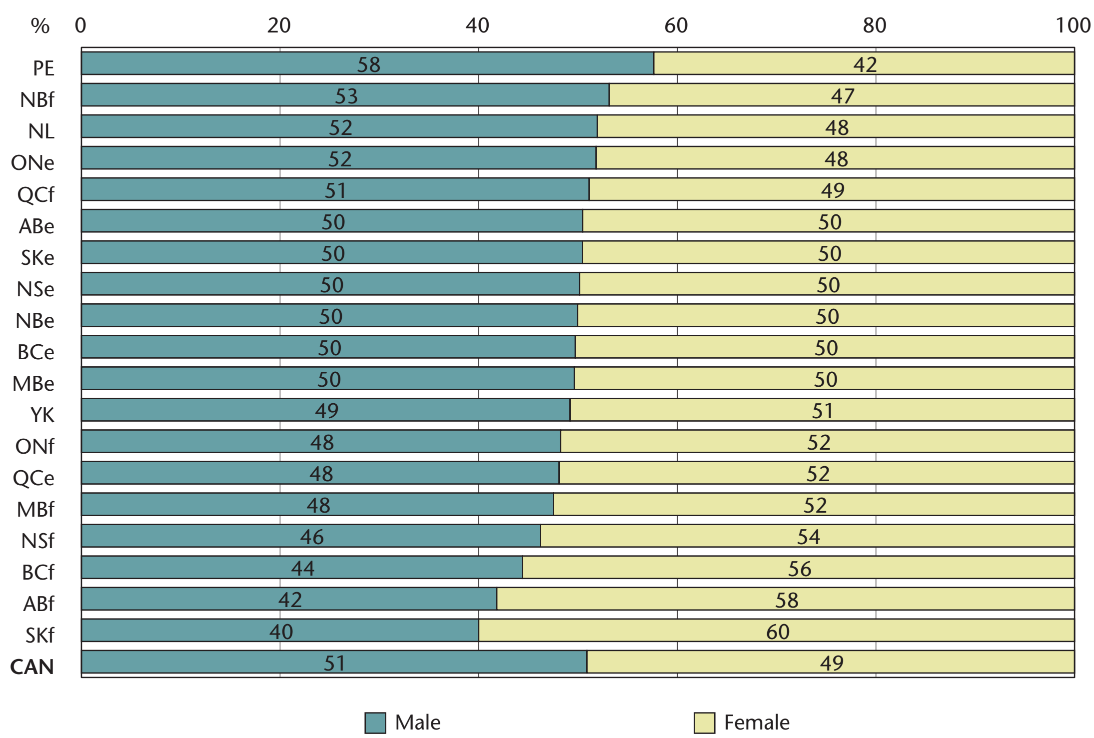
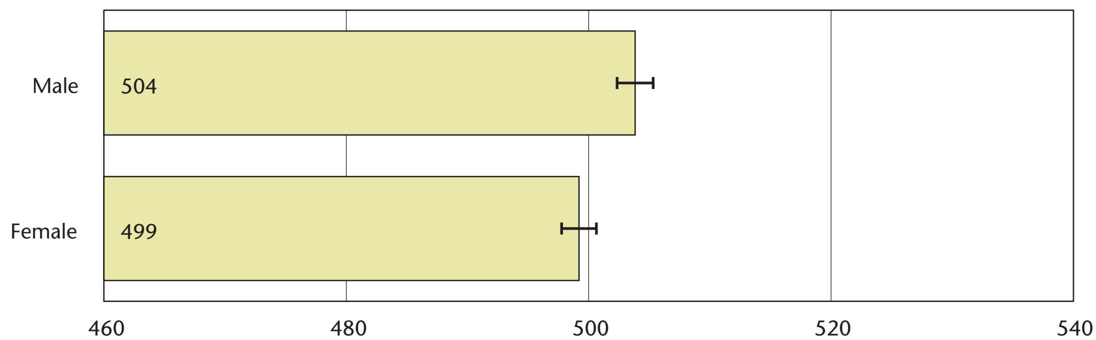


CHART 3-2 **Males and females by jurisdiction and language**



Gender differences in mathematics mean scores are shown in Chart 3-3. The comparison of means shows that males outperform females by a small but statistically significant margin. It is interesting to note that the differences in male and female participation rates in the test, as shown in Charts 3-1 and 3-2, may have had a small impact on the reported performance levels in jurisdictions for which these differences are found.

Chart 3-3 **Mean mathematics proficiency scores by gender**



Language

The PCAP populations were defined by the language of the school, and the tests were written in that language. However, the language of the school may not be the same as that used outside the school. Students were therefore asked to identify their mother tongue, defined as the language they first learned and still understand. Responses to this question by population are given in Chart 3-4.

Chart 3-5 shows mean mathematics scores for these language groupings. There is a statistically significant difference in mean scores between English and French populations and between students having an Aboriginal first language and all of the other groups. Those with a language other than one of the official languages or an Aboriginal language outperform English first language but not French first language students.

CHART 3-4 **First language by jurisdiction and language**[8]

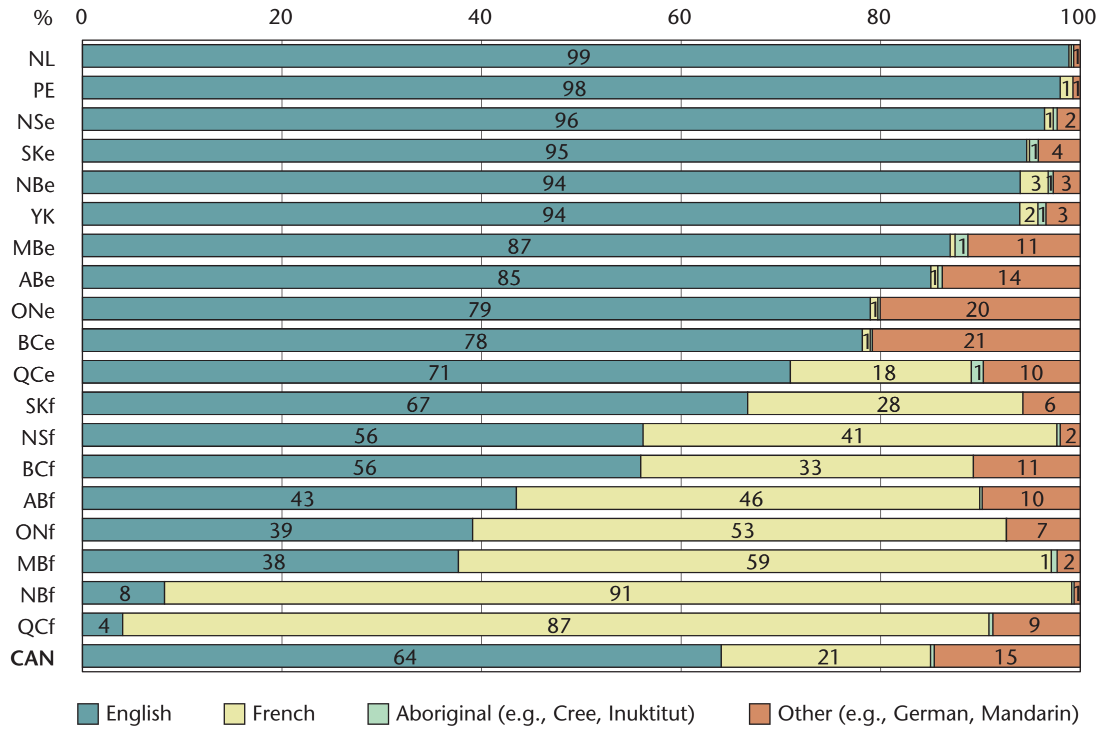
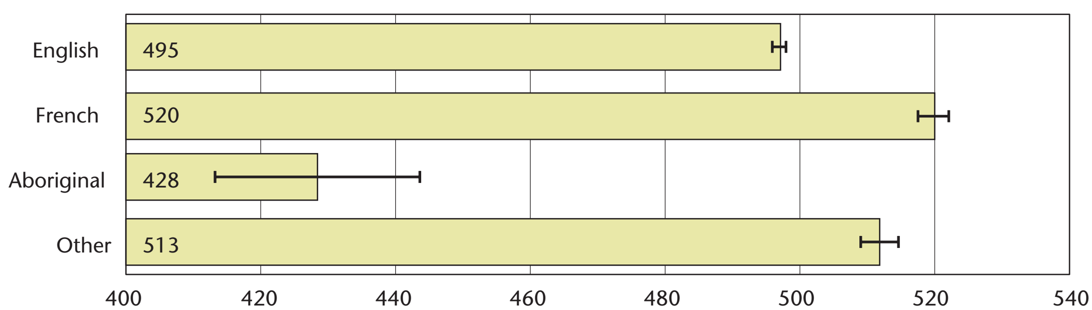


CHART 3-5 **Mean mathematics scores by first language**



[8] Throughout this report, where it is judged necessary to more clearly identify differences between jurisdictions, the results are sorted from highest to lowest on specific categories rather than following the conventional west-to-east order.

The relationship of language to achievement may be examined more closely by dividing the overall English and French populations into whether they represent the majority or the minority official language. Majority French and minority English are unique to Quebec while the other two groups are found in all of the other jurisdictions.

Chart 3-6 shows the proportions of each of these groups who speak each of the official languages in their everyday lives. This latter variable is a composite of responses to a series of questions on language used in various settings other than the school. The patterns for the two majority groups are similar, with most students speaking the same language as the language of the test. However, the results for the minority populations are quite different, with 43 per cent of minority English speaking French and 38 per cent of minority French speaking English in their everyday lives. This clearly illustrates the point that many minority language students are functioning in a language environment different from the language of the school.

Chart 3-6 **Main language used by majority and minority language groups in students' everyday lives**

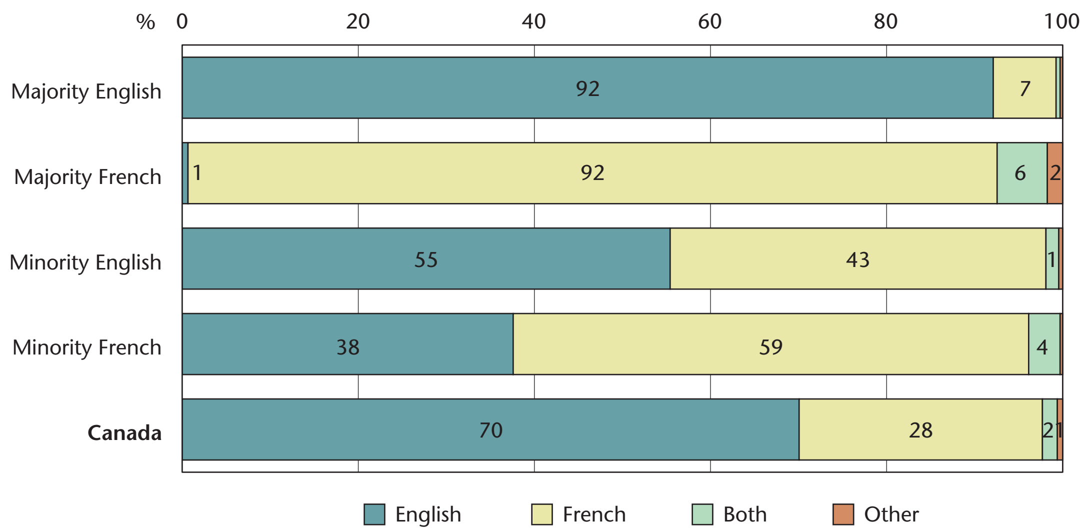
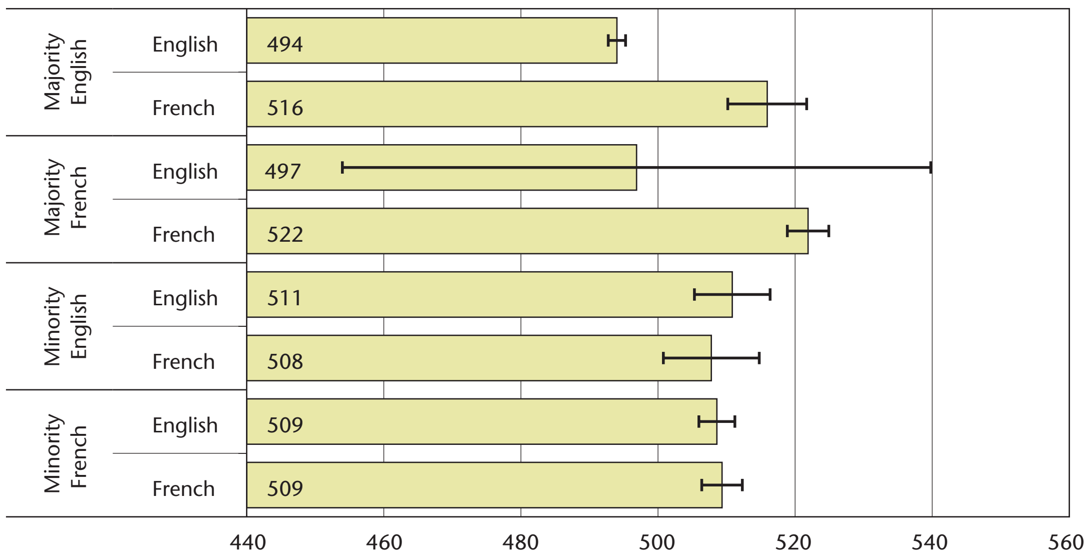


Chart 3-7 shows mean mathematics scores for four population groupings defined by language and by majority and minority status within the jurisdiction. The largest differences are those between French and English users in both majority English and majority French populations. However, the mean difference between the two majority French language sub-groups is not statistically significant because the small number of English speakers in this population results in a wide confidence interval, as shown by the error bar for that group. None of the other within-group differences are statistically significant.

Chart 3-7 **Mean mathematics scores by majority and minority language and main language used in students' everyday lives**



Student socioeconomic status

Two indicators of student socioeconomic status, mother's education and the number of books in the home, were included on the questionnaire. Charts 3-8 to 3-11 give results on these two indicators by jurisdiction and language. Differences across jurisdictions on these variables are relatively small. However, Chart 3-9 shows that, in general, francophone jurisdictions have lower numbers of mothers with less than high school and high school education than other populations. Chart 3-11 also shows that three francophone populations, New Brunswick, Quebec, and Ontario, report fewer books in the home than most other populations.

CHART 3-8 **Mother's education by jurisdiction**

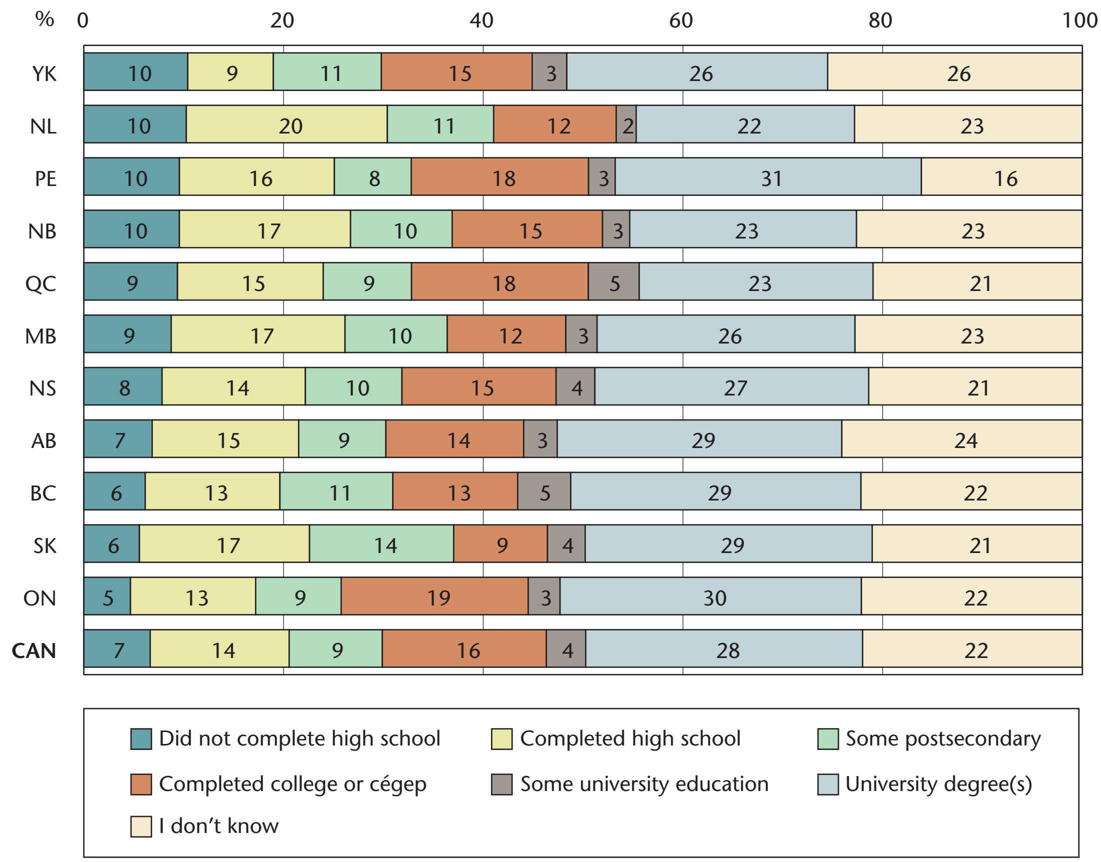


Chart 3-9 **Mother's education by jurisdiction and language**

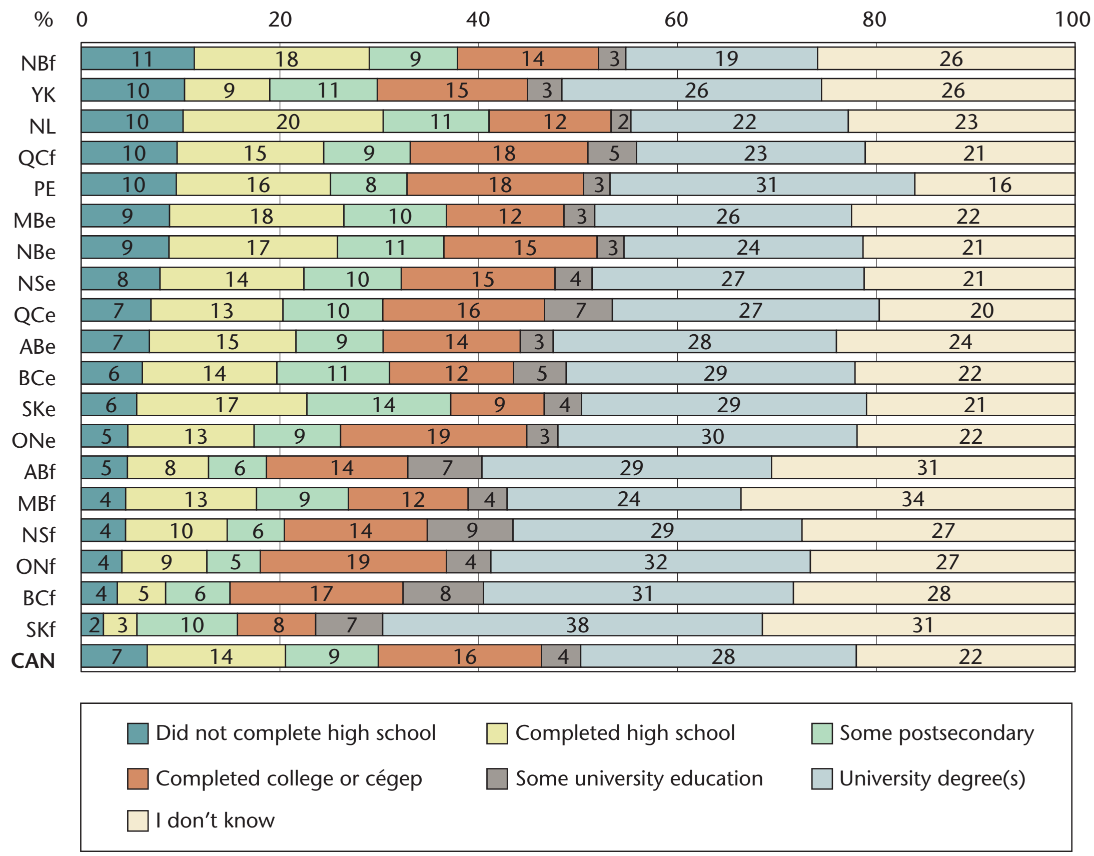


Chart 3-10 **Number of books in the home by jurisdiction**

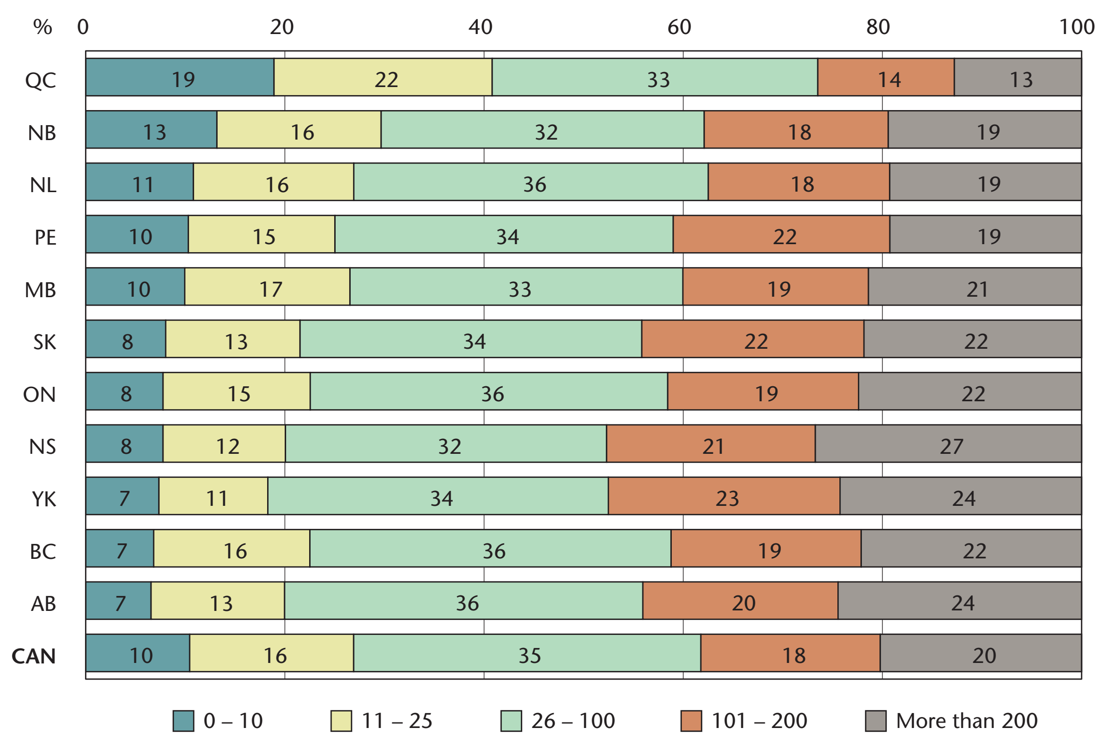
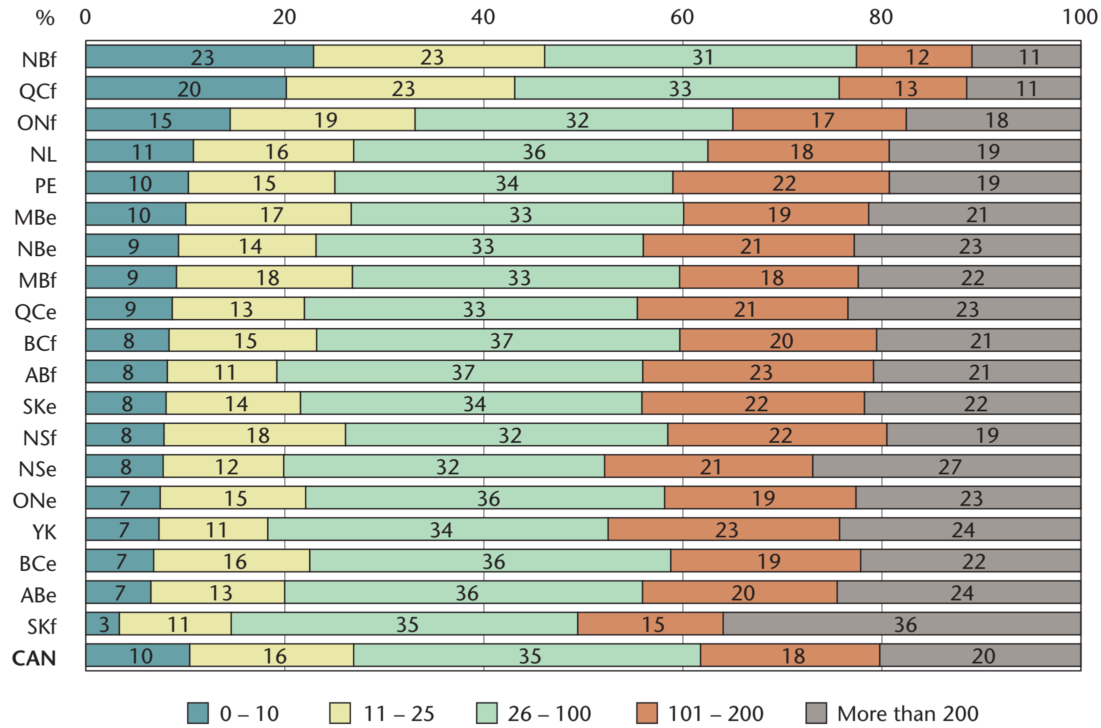


Chart 3-11 **Number of books in the home by jurisdiction and language**



Mean mathematics scores comparisons by these variables appear in Charts 3-12 and 3-13. The pattern is quite clear here. Having a mother with a higher level of education and having more books at home are both associated with higher performance.

Chart 3-12 **Mean mathematics scores by mother's education**

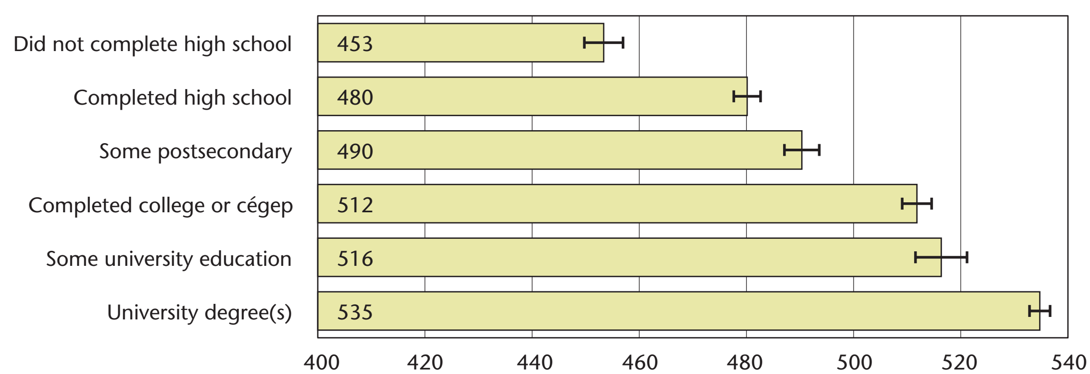
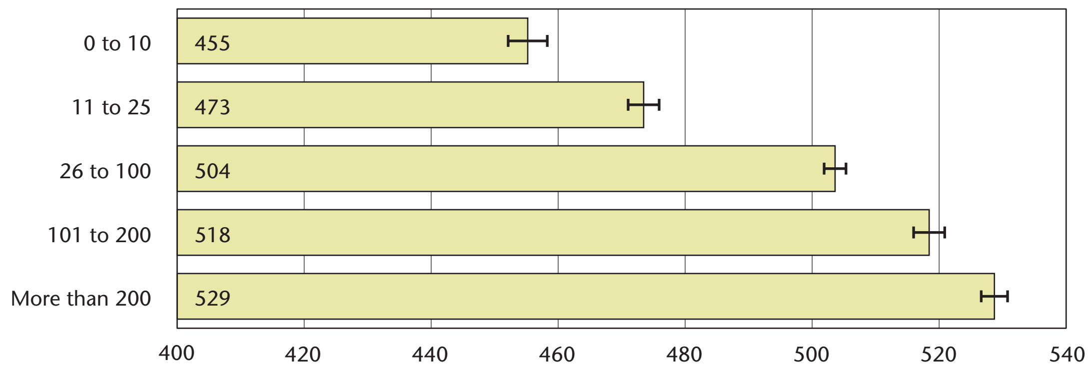


Chart 3-13 **Mean mathematics scores by number of books in home**



Immigration status

Charts 3-14 and 3-15 show that most students in all jurisdictions were born in Canada. The proportions born outside of Canada are quite variable across populations, ranging from 18 per cent in Ontario (English) down to 2 per cent or less in New Brunswick (French) and Newfoundland and Labrador.

The impact of immigration status on mathematics mean scores is shown in Chart 3-16. The difference in mean scores is statistically significant, with those born outside of Canada performing at higher levels.

Chart 3-14 **Percentage of students by whether or not born in Canada by jurisdiction**

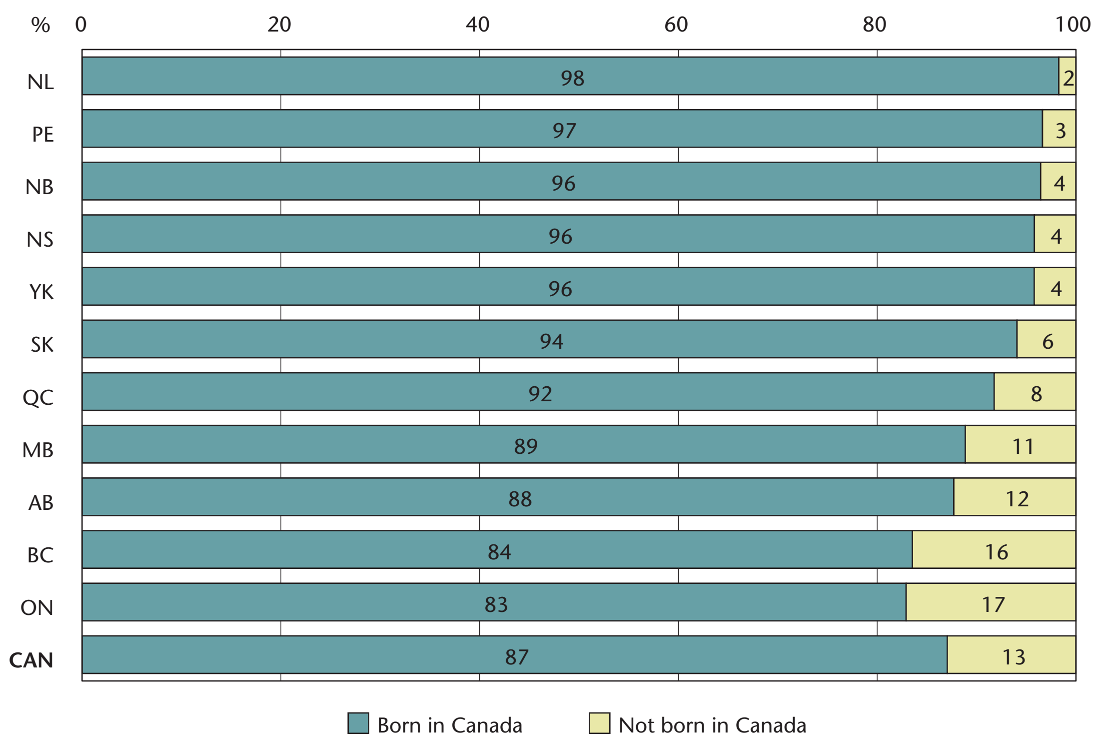


CHART 3-15 **Percentage of students by whether or not born in Canada by jurisdiction and language**

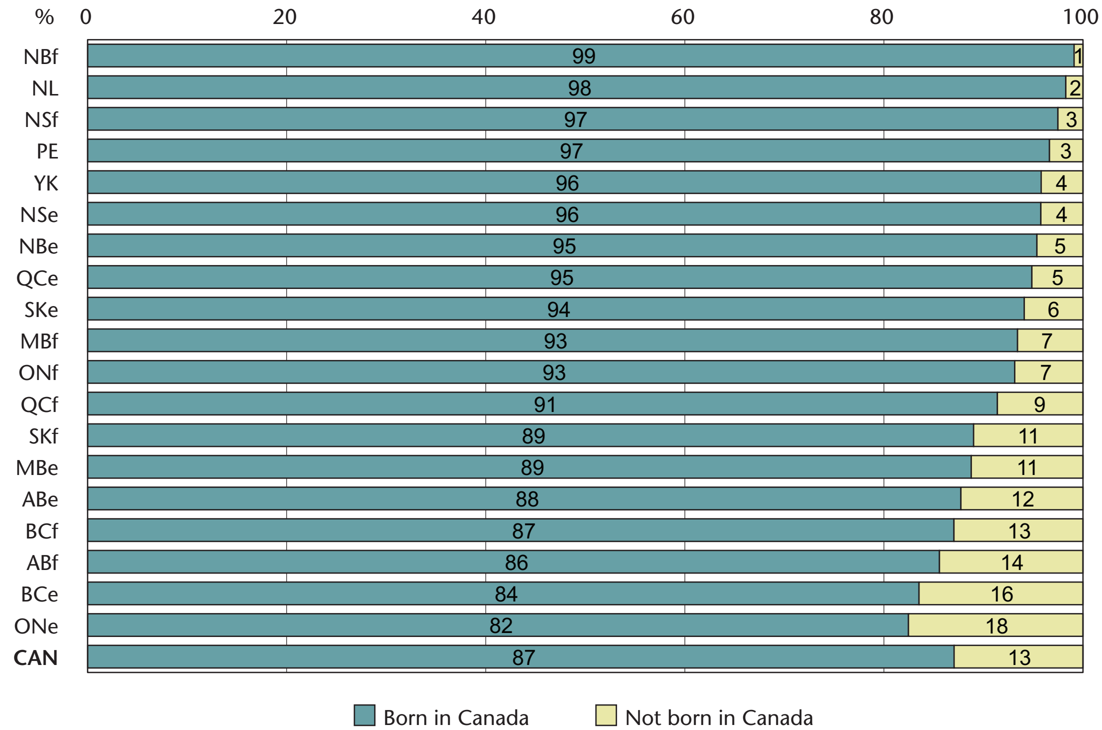
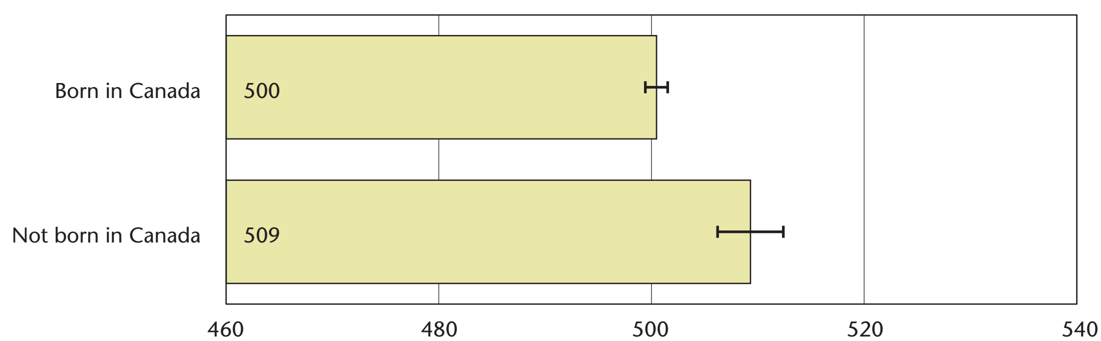


CHART 3-16 **Mean mathematics scores by whether or not born in Canada**



Aboriginal identity

Aboriginal identity was defined by the question "Are you of Aboriginal ancestry (for example, at least one of your parents/guardians is of Aboriginal ancestry)? *(Aboriginal ancestry refers to First Nations [North American Indian], Inuk [Inuit], and Métis.)*"

Charts 3-17 and 3-18 show that the percentage of students of Aboriginal identity in the public education system varies substantially by population, both overall and by specific Aboriginal groups. Generally, more Aboriginal students are found in Yukon and in the western provinces than in other jurisdictions.

The mathematics performance results in Chart 3-19 show that Aboriginal students have lower mathematics scores than non-Aboriginal students. However, Aboriginal students in Ontario performed significantly better than those in other jurisdictions which may be due in part to initiatives in Ontario to enhance numeracy and literacy. Among the Aboriginal groups, those of Métis identity perform significantly better than those of First Nations identity, although there remains much discussion amongst Aboriginal groups surrounding issues and definitions of identity. Nevertheless, the strategies and practises that promoted this higher achievement would benefit from further study.

CHART 3-17 **Percentage of students of non-Aboriginal and Aboriginal identity by jurisdiction**

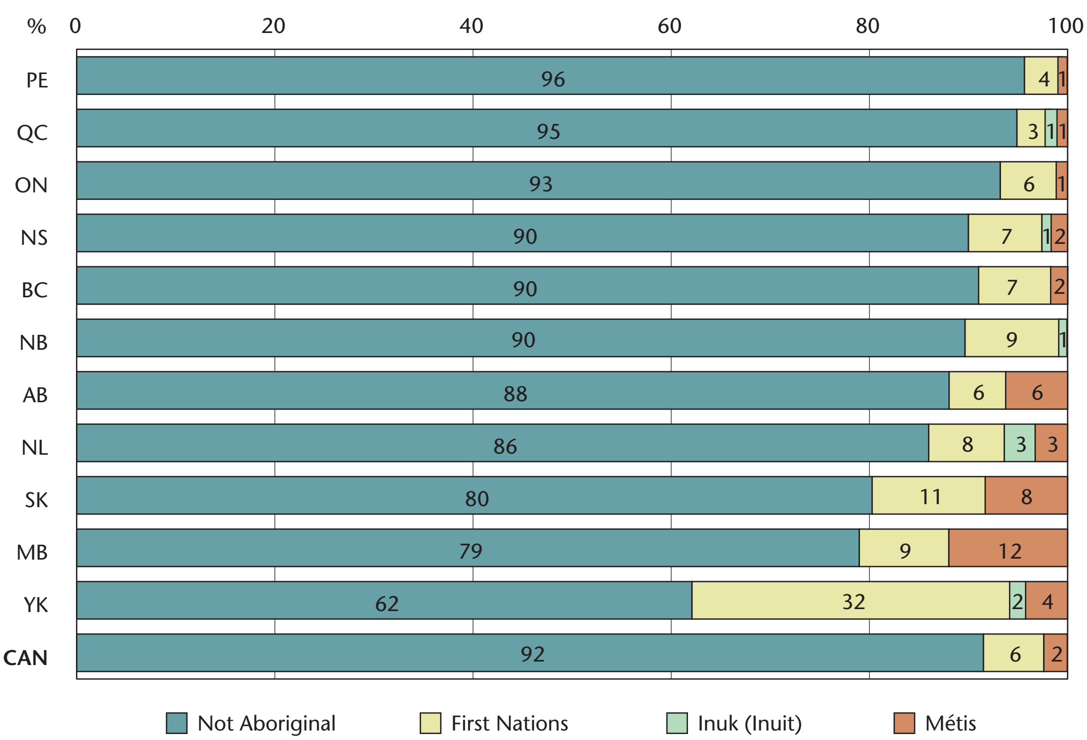


Chart 3-18 **Percentage of students of non-Aboriginal and Aboriginal identity by jurisdiction and language**

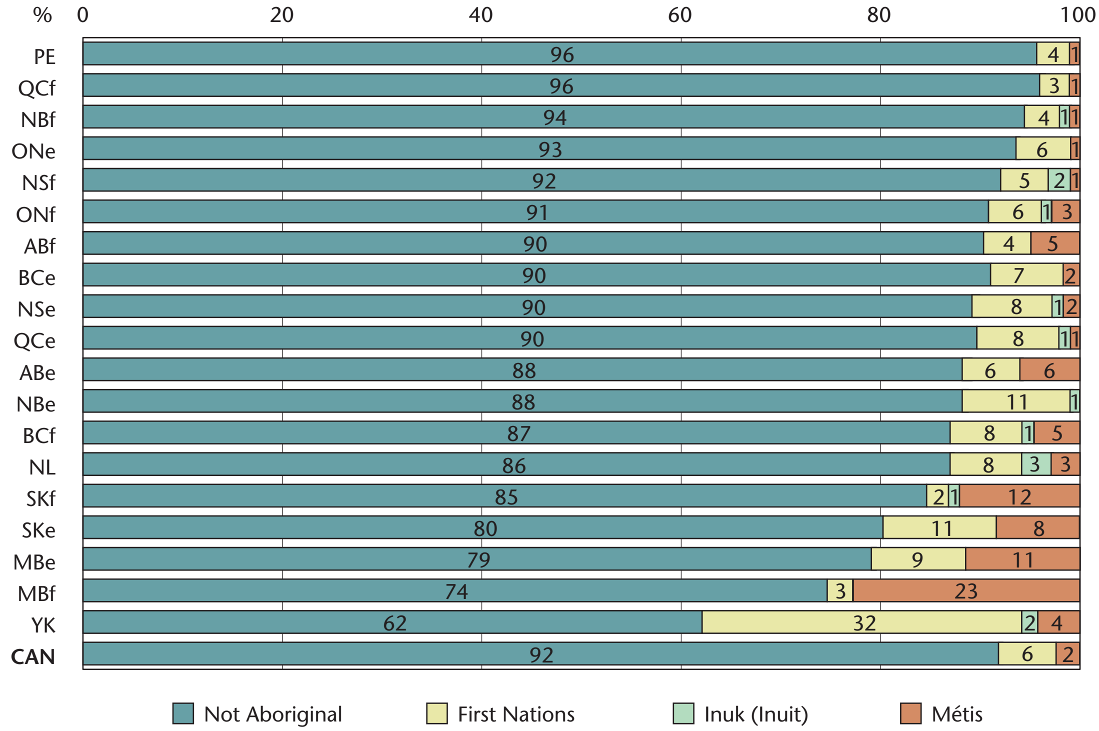
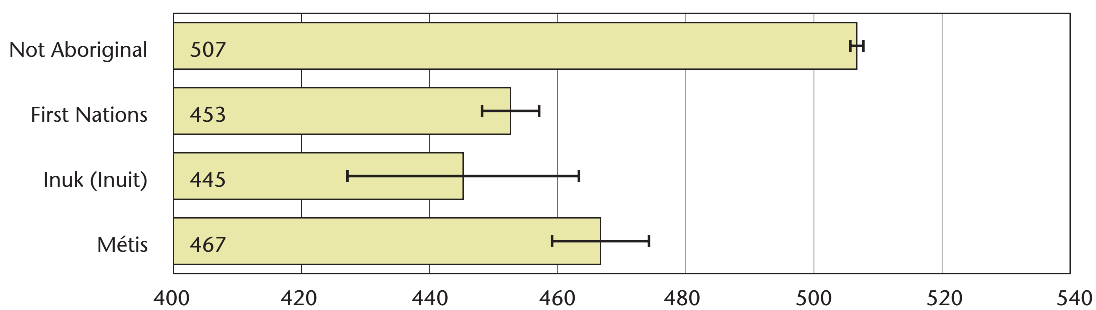


Chart 3-19 **Mean mathematics scores by non-Aboriginal and Aboriginal identity**



Student aspirations

Charts 3-20 and 3-21 show that Canadian students have high aspirations, with more than half aspiring to university graduation nationally and in most populations. A language division is apparent from Chart 3-21, with fewer francophone than anglophone students aspiring to university graduation. The number of "I don't know" responses is also larger for francophone students, suggesting that fewer francophone students have made up their minds about their future education at this stage of their school career.

Occupational aspirations, as shown by Charts 3-22 and 3-23, are also high, with more students aspiring to professional occupations that require university degrees than to any of the other occupational areas (examples of professions given in the question were nurse, lawyer, teacher). On the other hand, relatively few aspire to trades or business. Many reported that they don't know in what occupation they expect to find themselves, which likely relates to the age of the PCAP population.

Student aspirations are generally reflected in their relative mathematics performance, as shown in Charts 3-24 and 3-25. For education, those expecting only to complete high school have the lowest scores and those expecting to complete university the highest. In both cases, these extremes are significantly different from the scores of those in the other categories. The high school score is especially striking at more than 100 points below that for those aspiring to complete university. In the case of occupational aspirations, those expecting to work in professions or in information technology have significantly higher scores than those in other categories, while those who expect to work in trades or "other" areas have significantly lower scores.

CHART 3-20 **Student educational aspirations by jurisdiction**

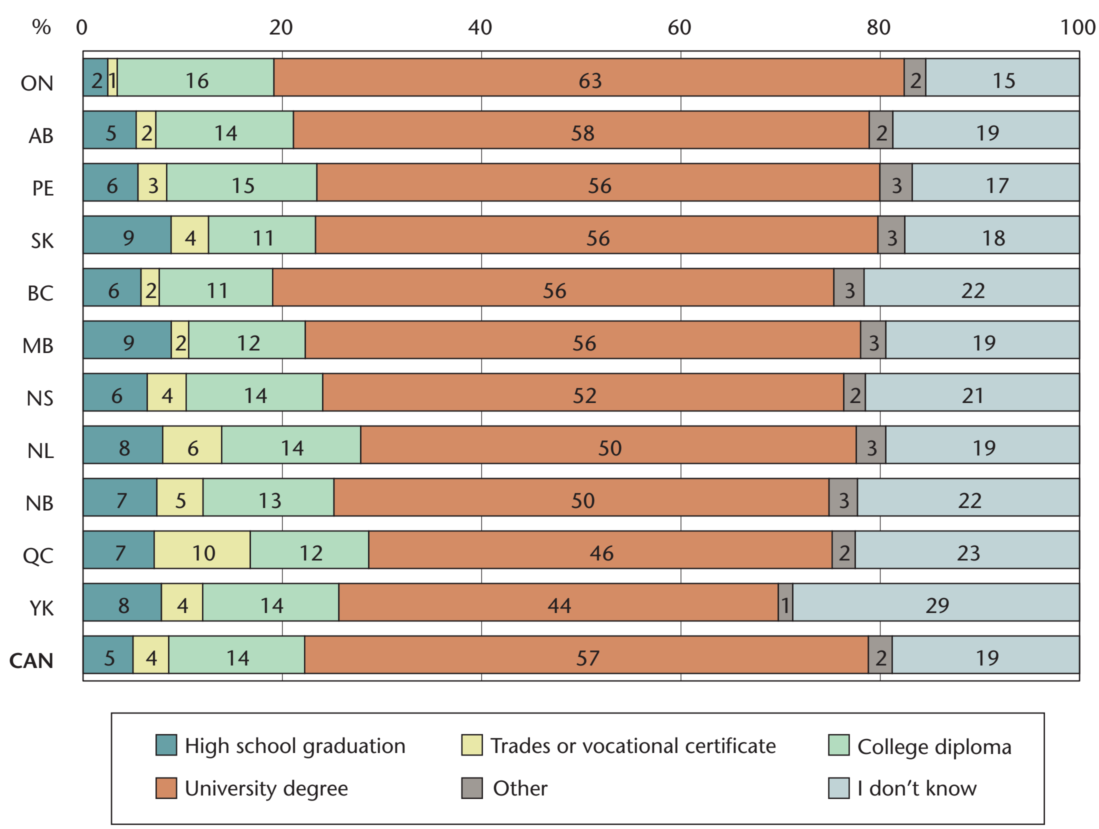


Chart 3-21 **Student educational aspirations by jurisdiction and language**

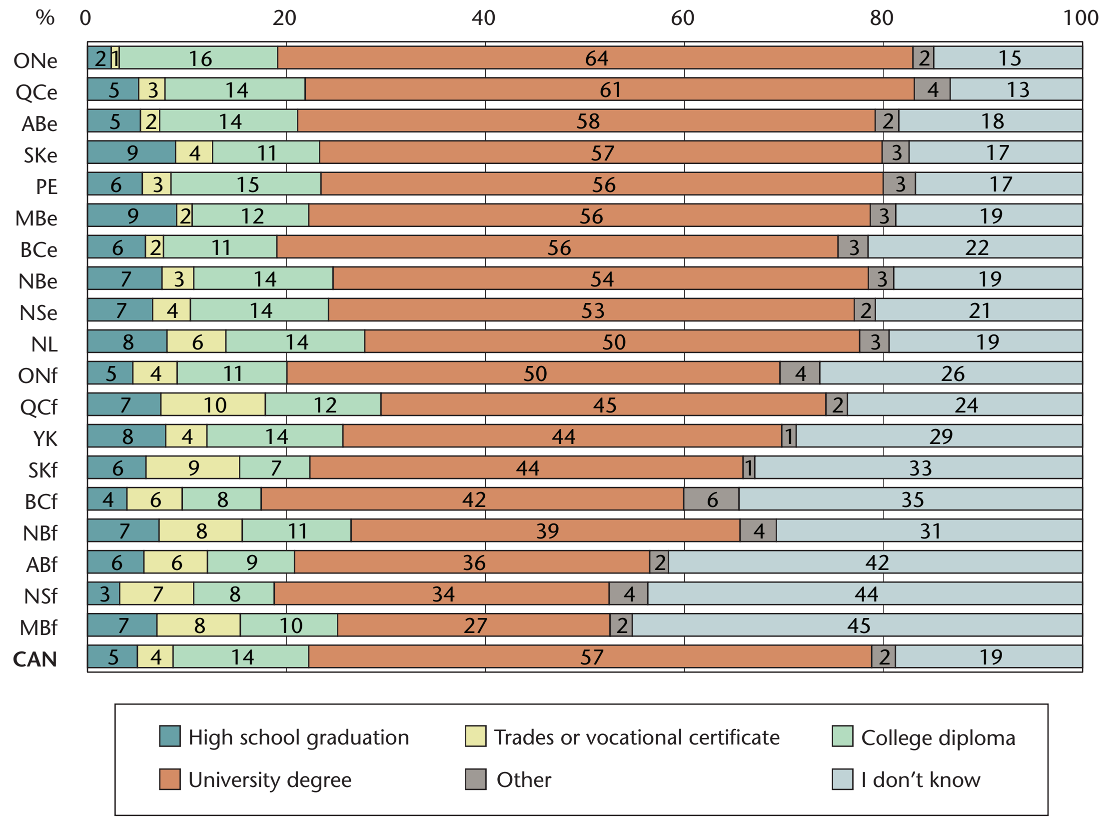


Chart 3-22 **Student occupational aspirations by jurisdiction**

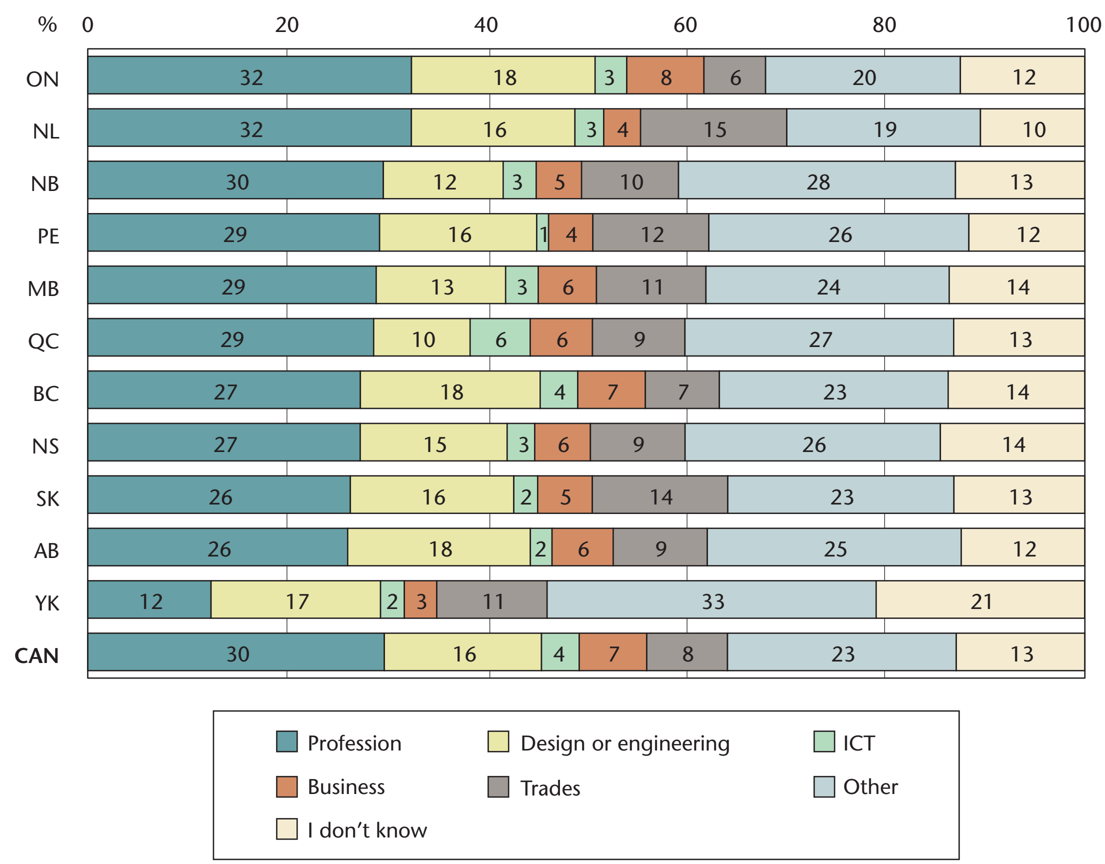


CHART 3-23 **Student occupational aspirations by jurisdiction and language**

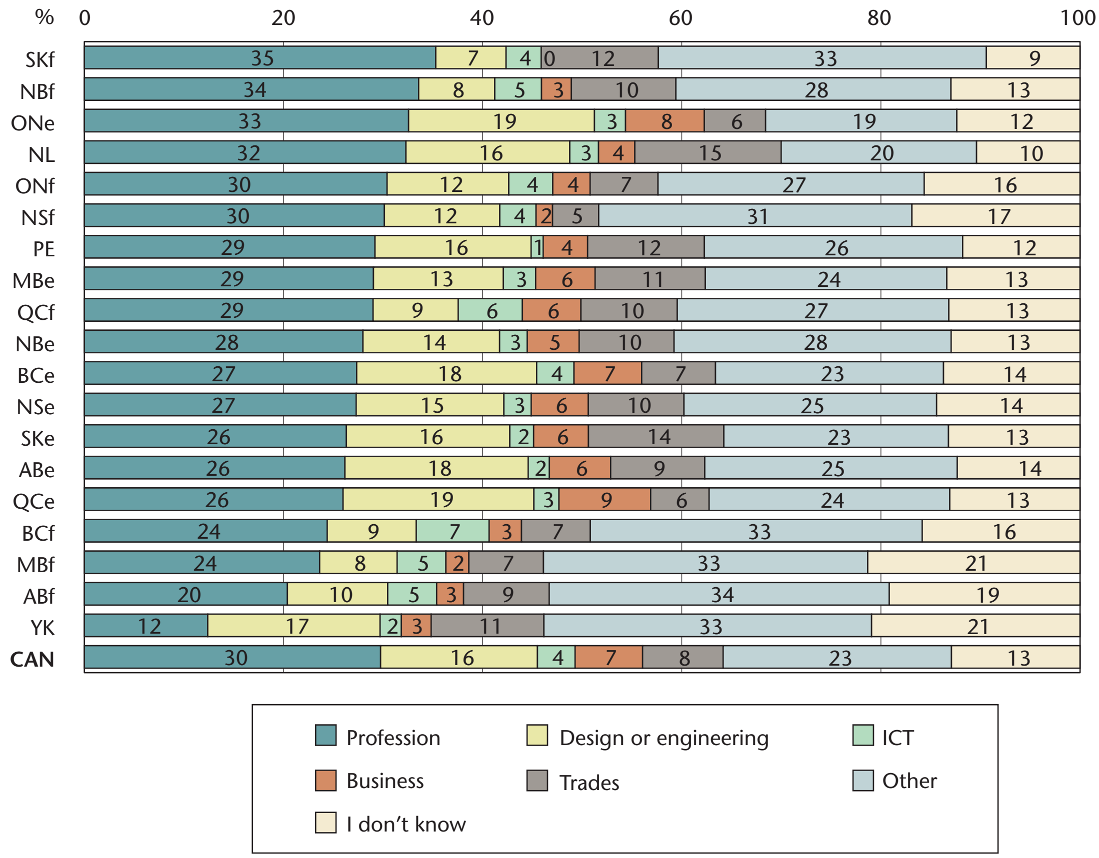


CHART 3-24 **Mean mathematics scores by educational aspirations**

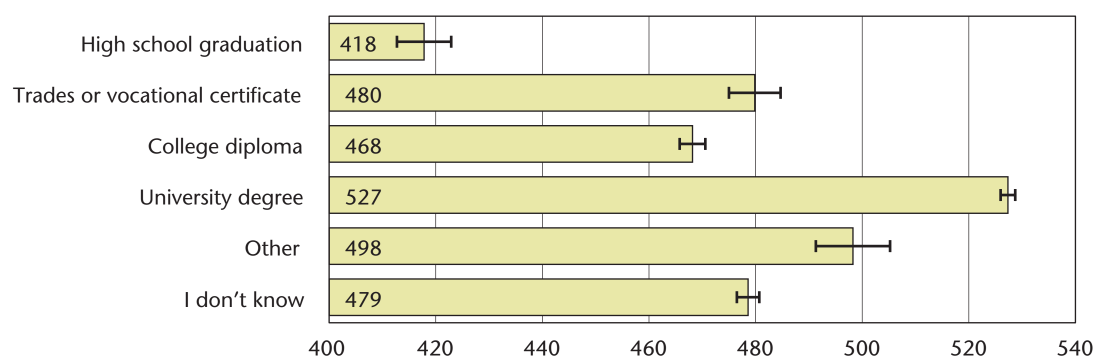
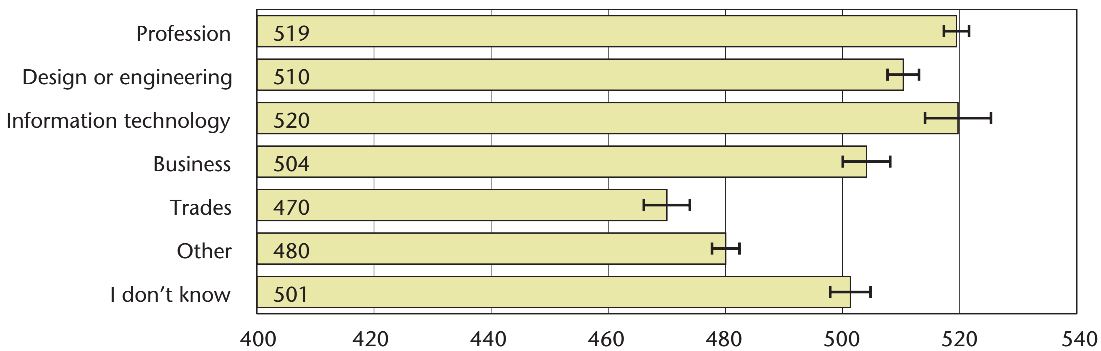


Chart 3-25 **Mean mathematics scores by occupational aspirations**



Teacher characteristics

Teacher gender

Chart 3-26 shows that overall there are slightly more female than male mathematics teachers. Most populations are close to the Canada mean. However, there are some extremes, with Manitoba French having proportionally the most male teachers and New Brunswick French and Yukon the fewest.

Chart 3-27 shows no significant difference in the mean mathematics scores of students taught by male and female teachers.

Chart 3-26 **Male and female teachers by jurisdiction and language**

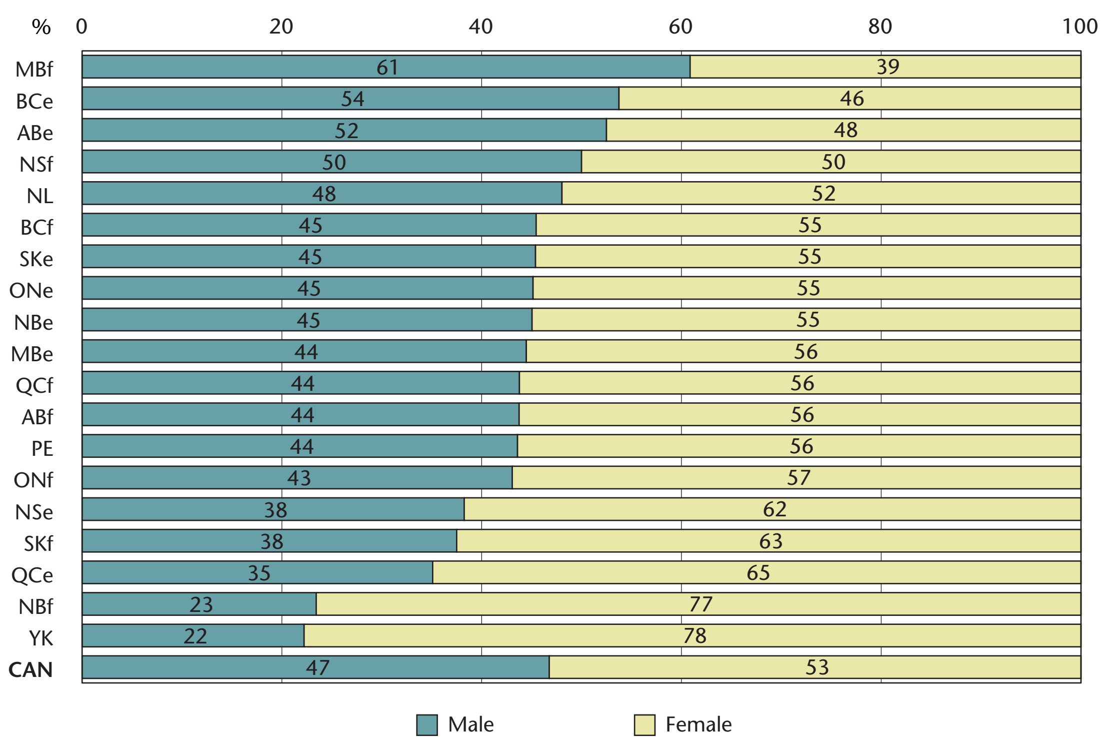
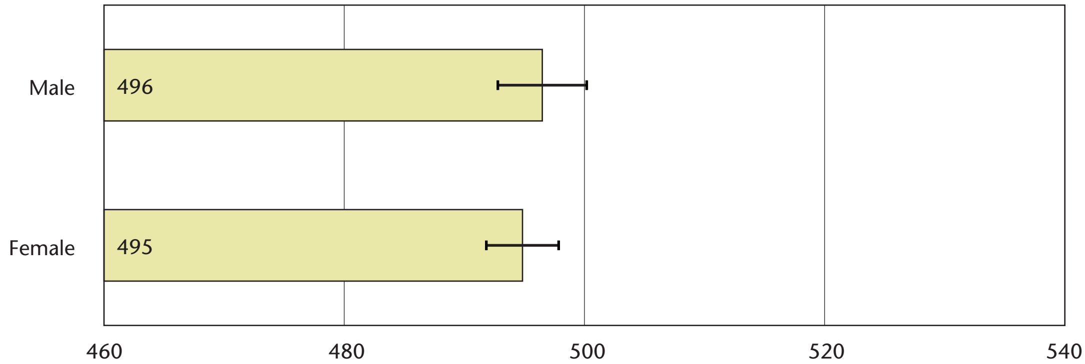


CHART 3-27 **Mean teacher mathematics scores by teacher gender**[9]

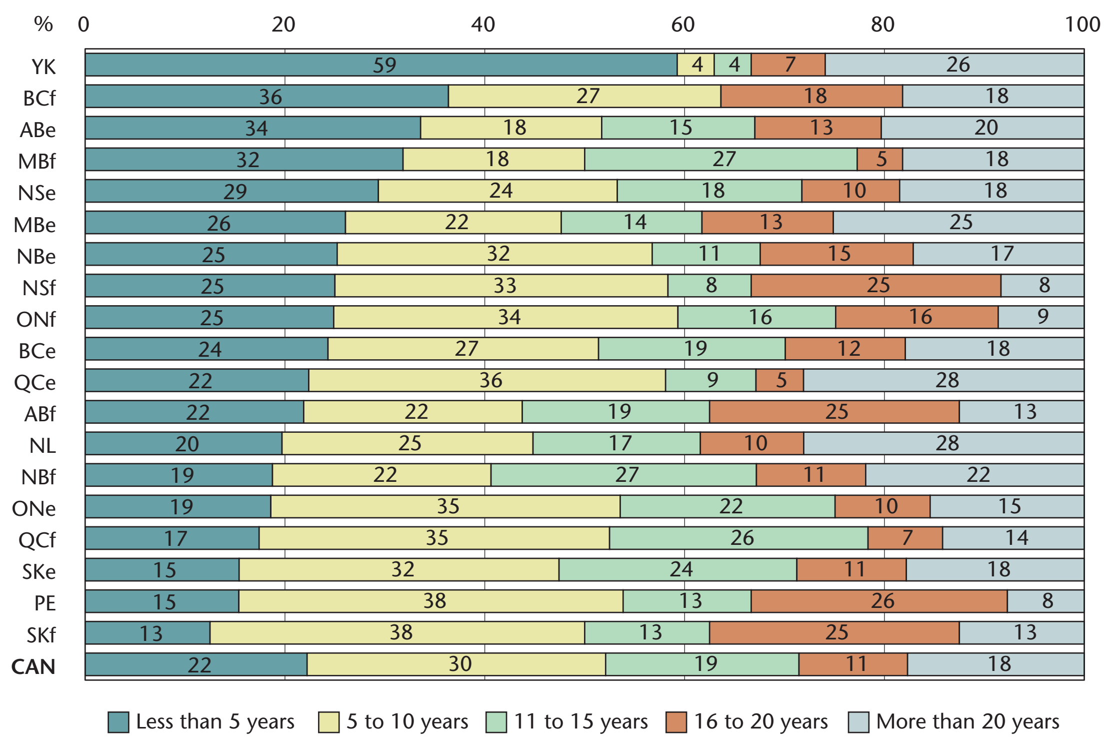


Teaching experience

Chart 3-28 shows a broad range of teacher experience in all populations except Yukon, where a majority of teachers have less than five years' experience. The proportion of teachers with less than five years' and five to ten years' experience is generally higher than the proportions in older age groups, particularly those with more than 20 years' experience, even though the latter represents a wider range of years.

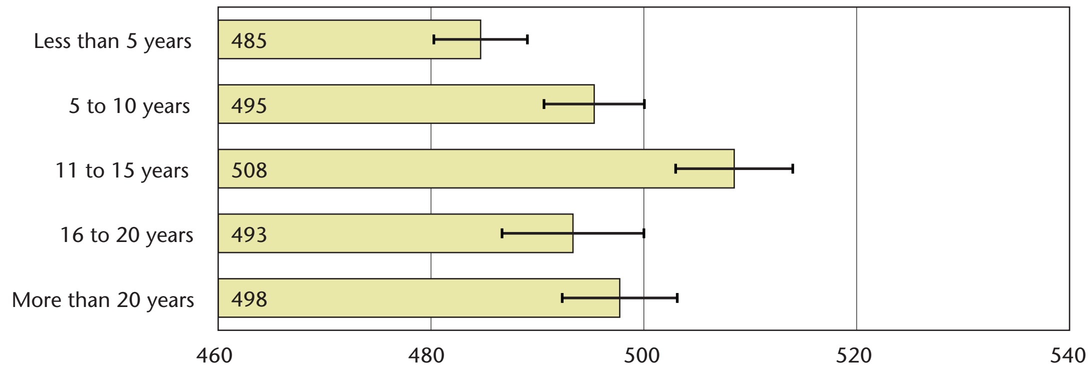
Chart 3-29 shows a non-linear pattern, with an increase in mathematics performance with experience up to the 11–15-year range and a decline thereafter.

CHART 3-28 **Range of teaching experience by jurisdiction and language**



[9] The basic unit used to compute each mean in teacher level charts is the mean over all students taught by a teacher. The reference to "mean teacher mathematics scores" reflects the fact that these are "means of means" and are unweighted with respect to the number of students taught by a teacher. At the population level, these are different from the means computed over all students because the number of students taught by a teacher differs across teachers. The same principle applies to school-level charts, where the basic unit is the mean for all students in a school.

Chart 3-29 **Mean teacher mathematics scores by teacher experience**



Teacher qualifications and specialization in mathematics

In most jurisdictions, there are strong incentives for teachers to improve their qualifications, particularly through acquiring additional university degree credentials. Chart 3-30 shows wide variation in the degree combinations held.[10] While most teachers (81 per cent for the country as a whole) hold a B.Ed. degree, the proportion holding a B.Ed. as the only degree varies widely across populations. In about half the populations, a majority of teachers hold more than one undergraduate degree. However, this is also widely variable. In most jurisdictions, those holding master's degrees or equivalent account for 20 per cent or less. Newfoundland and Labrador, Prince Edward Island, and British Columbia French each have more than 30 per cent of teachers with master's degrees.

Chart 3-31 shows that there are only small differences in mathematics achievement across various teacher degree combinations. The exception is the B.Ed. only, for which achievement is significantly lower than any of the other combinations. An important point here is that holding a master's degree seems to convey no achievement advantage, despite the incentives in place for teachers to take this degree.

[10] The numbers in Chart 3-30 should be treated with greater caution than for other graphs, especially for small populations, because teachers could check more than one category; missing data were not separately coded for the degree categories; and the categories overlap to some extent. The numbers given therefore do not sum to 100 per cent.

Chart 3-30 **Teacher undergraduate university degrees by jurisdiction and language**

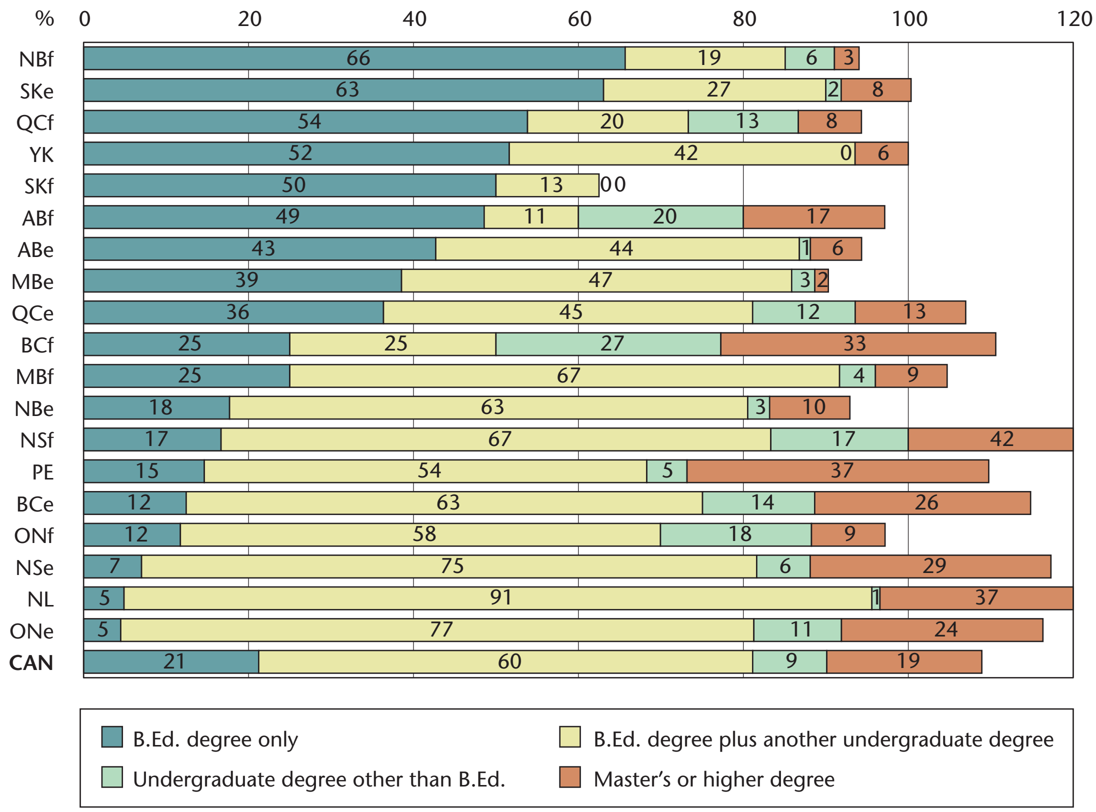
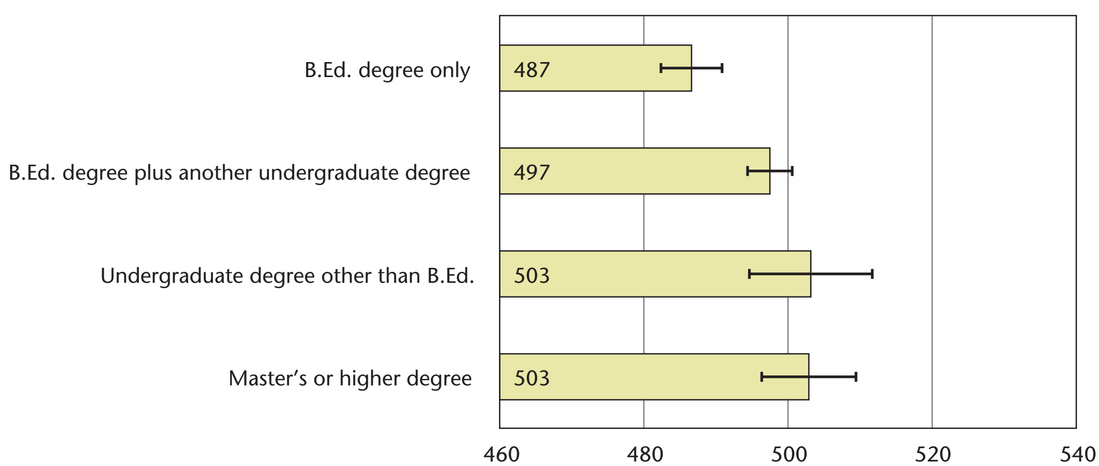


Chart 3-31 **Mean teacher mathematics scores by teacher university degrees**



Teacher specialization in mathematics was measured by two questions, the number of mathematics courses completed and the proportion of the teacher's assignment that is in mathematics. These results are given in Charts 3-32 and 3-33. Both of these variables show wide variations across populations. By the mathematics course measure, Quebec French and Newfoundland and Labrador have the most highly specialized teachers, with 47 per cent having 10 or more mathematics courses. At the opposite extreme are Ontario English, Manitoba French and English, and Saskatchewan French with 5 per cent or fewer teachers having that level of mathematics specialization. The two Quebec populations have the most specialized teachers in terms of teaching assignments. The patterns are fairly highly correlated, with teachers who are more specialized in mathematics by training also tending to have the greatest part of their teaching assignment devoted to mathematics.

Charts 3-34 and 3-35 show the relationship of mathematics specialization to mathematics achievement. The number of semester courses taken by teachers in mathematics shows no significant relationship with achievement. For the assignment variable, those reporting more than 70 per cent of their assignment to be in mathematics show higher achievement than those with smaller proportions.

Chart 3-32 **Number of mathematics semester courses taken by teachers by jurisdiction and language**

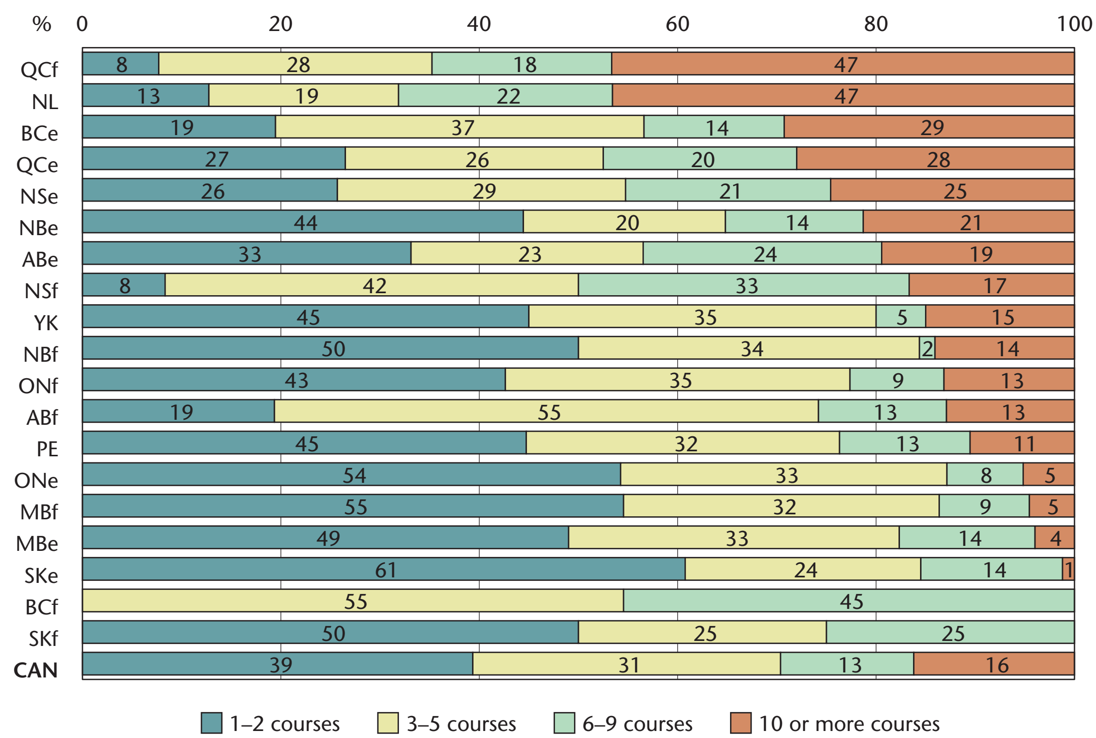


Chart 3-33 **Proportion of teachers specializing in mathematics by teaching assignment by jurisdiction and language**

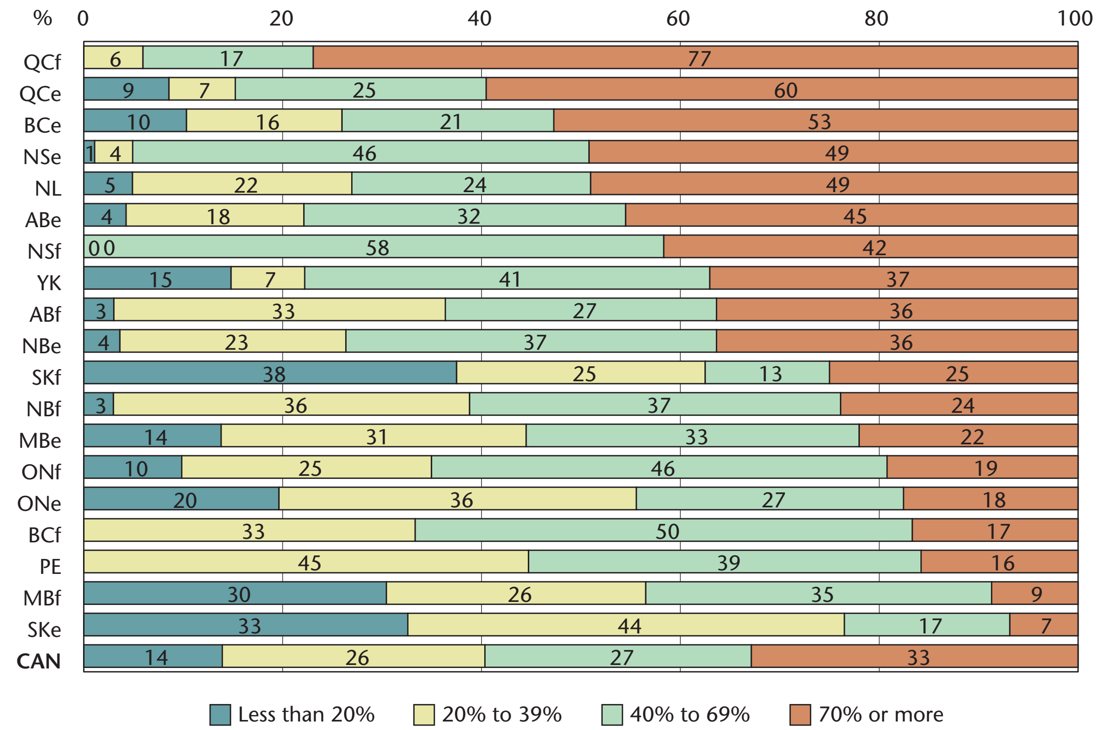


Chart 3-34 **Mean mathematics scores by semester courses taken by teacher in mathematics**

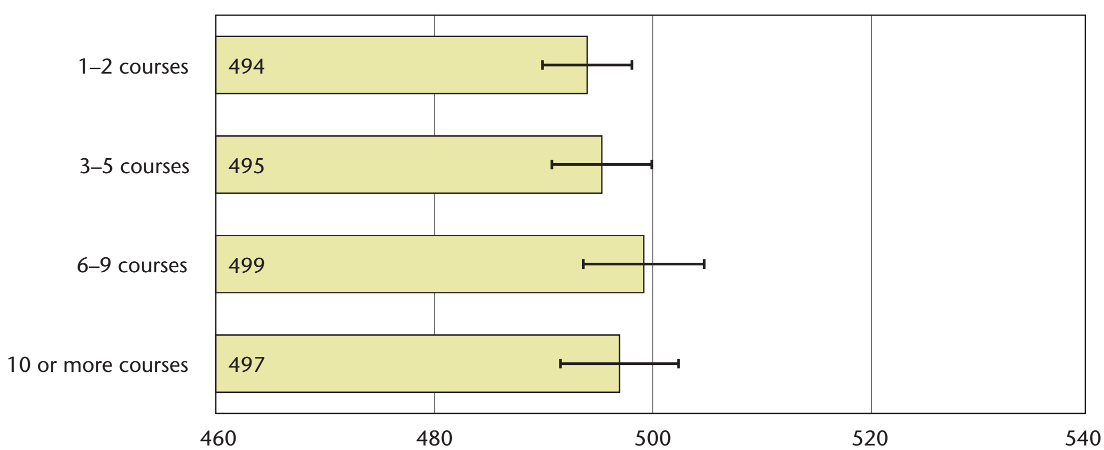
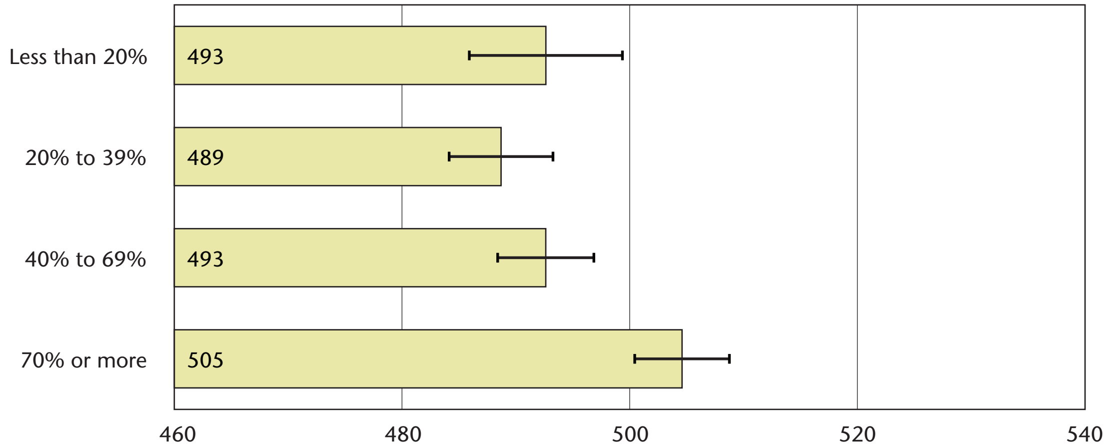


CHART 3-35 **Mean mathematics scores by percentage of teaching assignment in mathematics**



A third question in this sequence asked teachers to report the number of days of professional development in mathematics they had participated in over the past five years. These results are shown in Chart 3-36. Again, the pattern is one of wide variation in participation both within and across jurisdictions. Chart 3-37 indicates that the number of days of professional development in mathematics has no effect on mathematics achievement.

CHART 3-36 **Days of mathematics professional development in the past five years by jurisdiction and language**

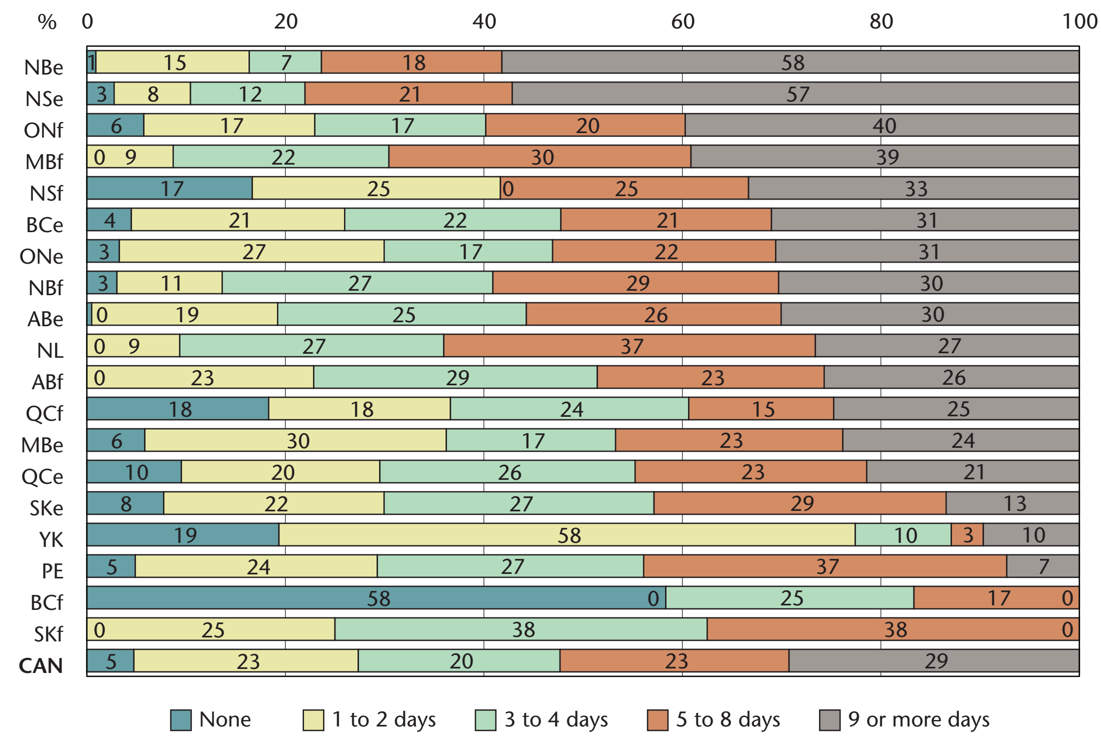
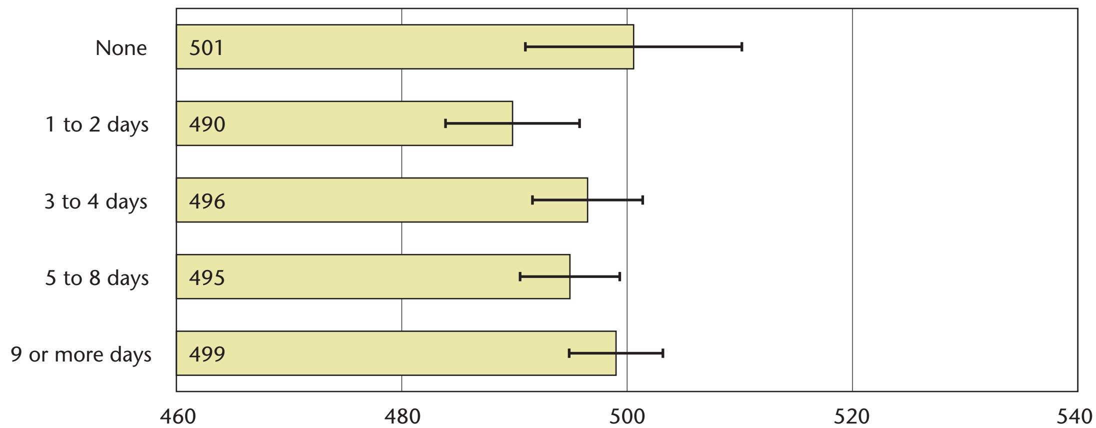


Chart 3-37 **Mean mathematics scores by number of professional development days in mathematics**



School characteristics

School size

Chart 3-38 shows the distributions of school size by population. It is clear that school size varies widely both within and across jurisdictions. Saskatchewan French and Yukon have the highest proportion of small schools, with enrolment less than 100, while Quebec French has the largest proportion of schools with total enrolment greater than 1,000.

Chart 3-39 shows the effect of enrolment on mathematics achievement. Schools with enrolments greater than 500 have higher achievement than those with enrolments in the 100–500 range. Schools with enrolments 100 or less show no significant difference from other size ranges. However, the margin of error is quite large for this category.

Chart 3-38 **Total school enrolment by jurisdiction and language**

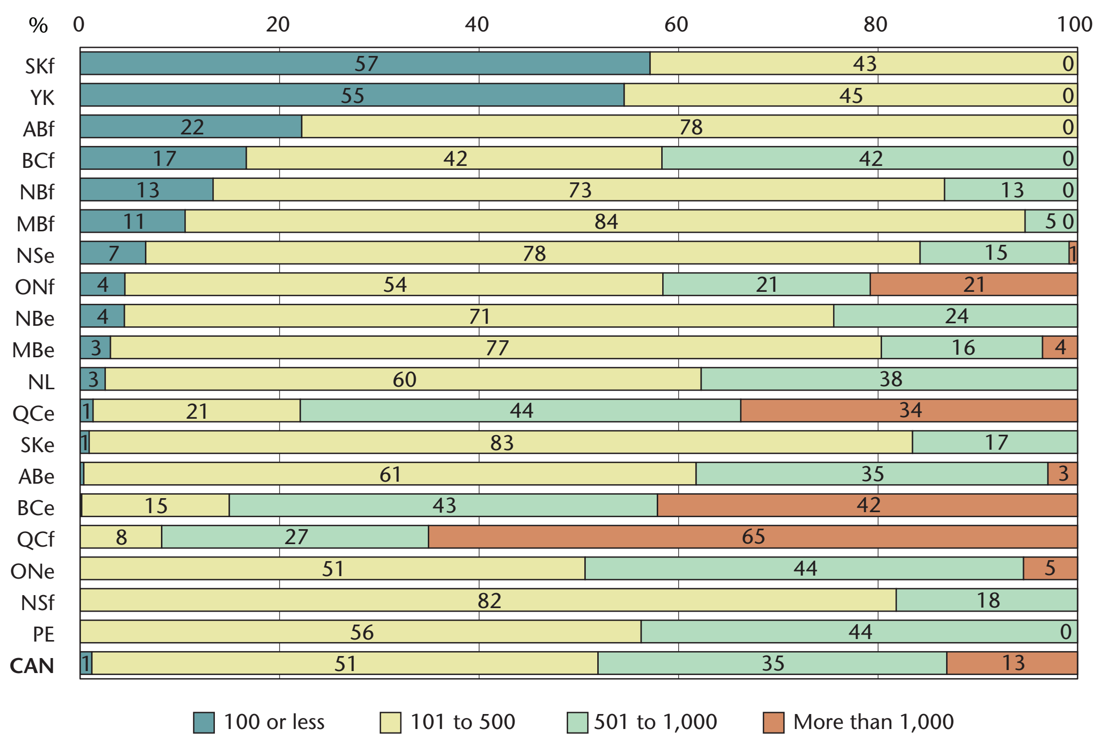
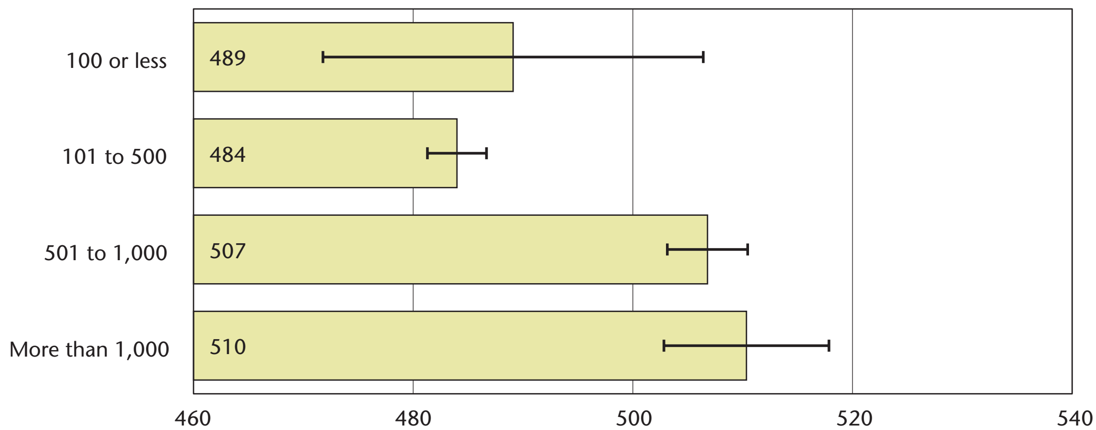


Chart 3-39 **Mean school mathematics scores by school enrolment**



Public and private schools

Chart 3-40 shows the percentages of schools identified by their principals as public or private. These were defined by whether the school is governed by a public (e.g., a public school board or similar authority) or a private body (such as a religious organization or a business). It is evident that the number of private schools is very small in most jurisdictions. The notable exceptions are Quebec, both English and French, and Manitoba English with percentages of 10 per cent or more.

Mean mathematics achievement by school governing structure is shown in Chart 3-41. It is clear that students in private schools significantly outperform those in public schools. This finding is of interest because the proportion of private schools is high enough in a few jurisdictions to influence the overall results for the jurisdiction.

A common argument for high performance on the part of private school students is that many of these students come from higher socioeconomic status families. It is therefore possible that the observed results would change if socioeconomic status were controlled. The models to be presented later in this chapter will shed some light on this issue.

Chart 3-40 **Percentages of public and private schools by jurisdiction and language**

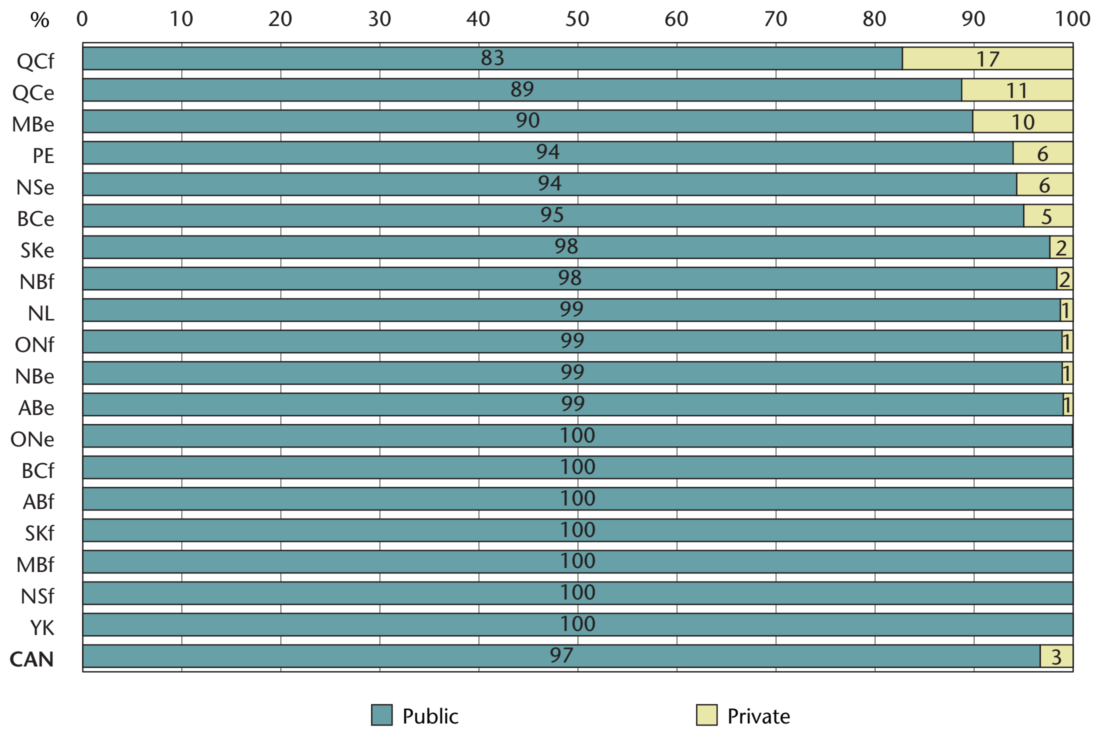
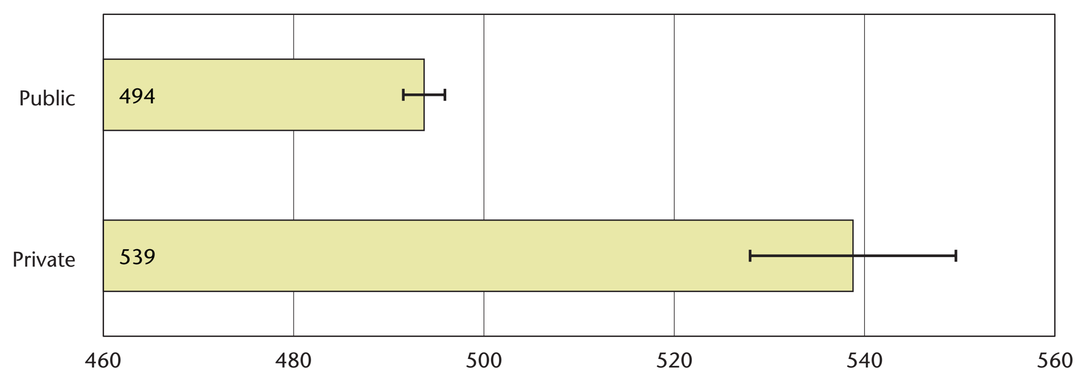


Chart 3-41 **Mean school mathematics scores by school governance**



Diversity of student populations

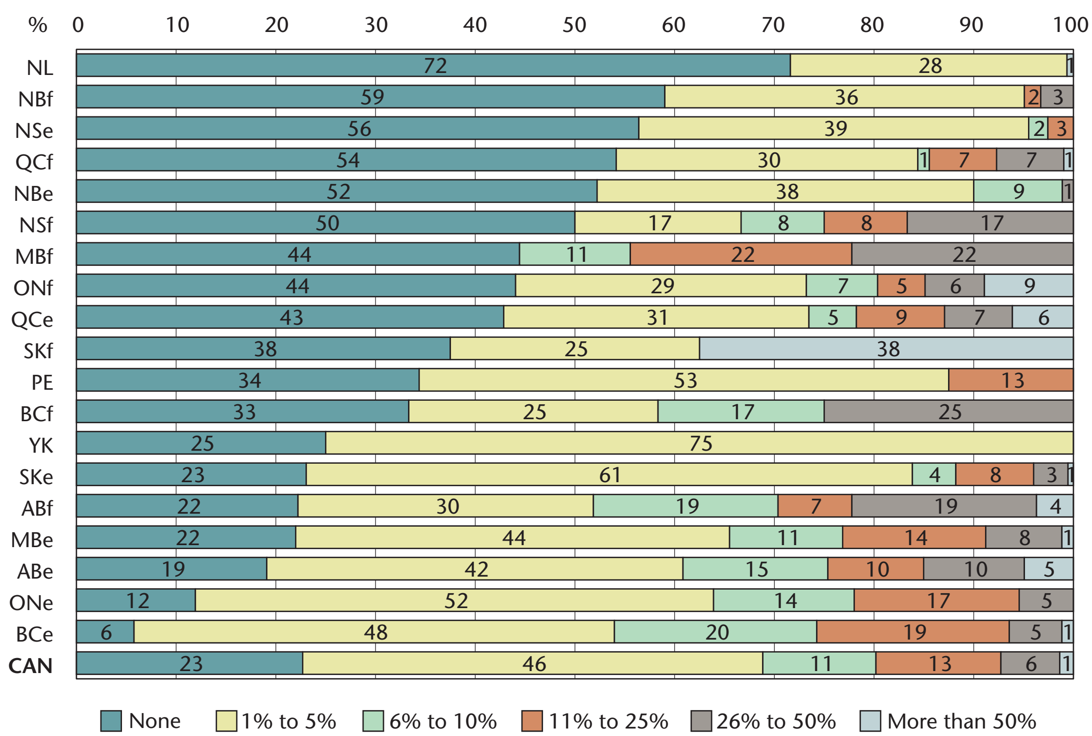
Two indicators of the diversity of school populations, the proportion of students in English or French second language (ESL/FSL)[11] programs and the proportion of students of Aboriginal identity in the school, were included in the school questionnaire. Distributions for these two variables are given in Charts 3-42 and 3-43.

Most of the francophone populations other than Quebec and New Brunswick stand out as having larger proportions than others of students in ESL/FSL programs. While such programs are usually associated with immigrant students, it is also possible that these programs apply to students who are in a school with a different official language, particularly French, from the language of the home.

The proportions of students of Aboriginal identity in the publically funded schools of most jurisdictions are relatively small. However, there are notable exceptions, with some 20 per cent or more having more than 25 per cent Aboriginal students in Yukon, Manitoba English and French, and Saskatchewan English.

Mean mathematics scores for schools with various proportions of students in these two categories are given in Charts 3-44 and 3-45. The pattern for ESL/FSL is non-linear, with schools in the 1 per cent to 5 per cent range of such students having the highest achievement and those in the 26 per cent to 50 per cent range the lowest. The pattern for Aboriginal students is more linear. Further study is needed to determine what factors lead to higher achievement in some jurisdictions with high student diversity and to identify how these students might be better served in the public education system.

CHART 3-42 **Percentages of schools with ESL/FSL students by jurisdiction and language**



[11] The terms ESL and FSL refer to students whose first language is different from the language of the school. Many ESL/FSL students are from immigrant families, but some are also from Canadian families who send their children to schools in the official language other than their home language.

CHART 3-43 **Percentages of publically funded schools with students of Aboriginal identity by jurisdiction and language**

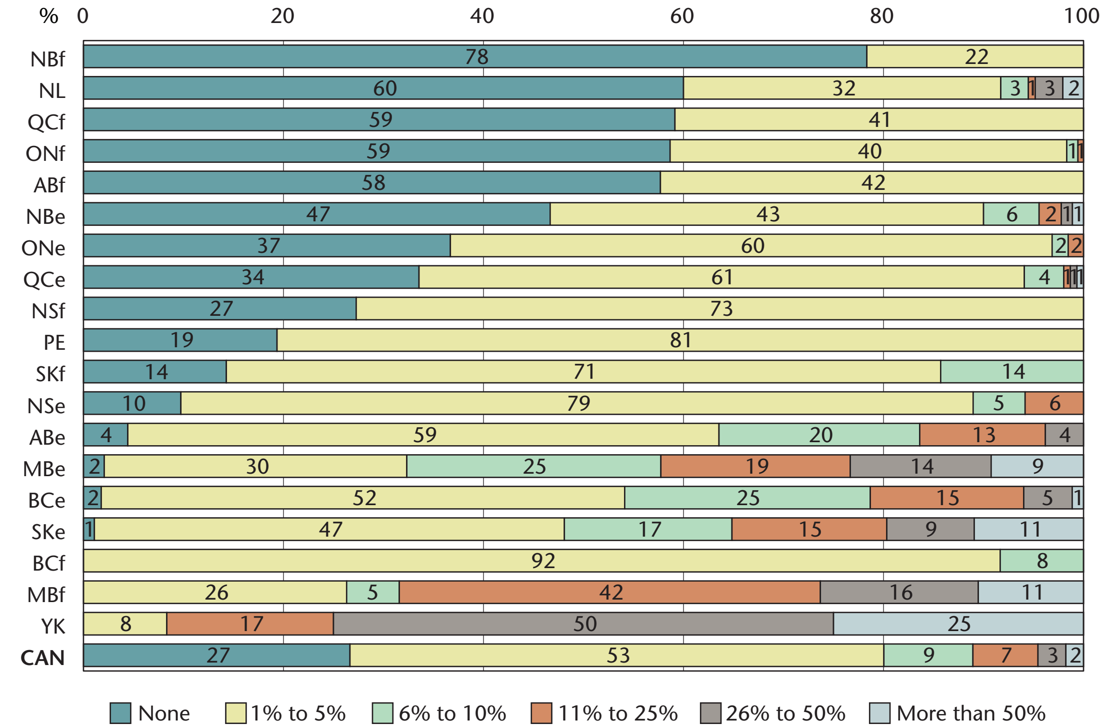


CHART 3-44 **Mean school mathematics scores by proportions of ESL/FSL students**

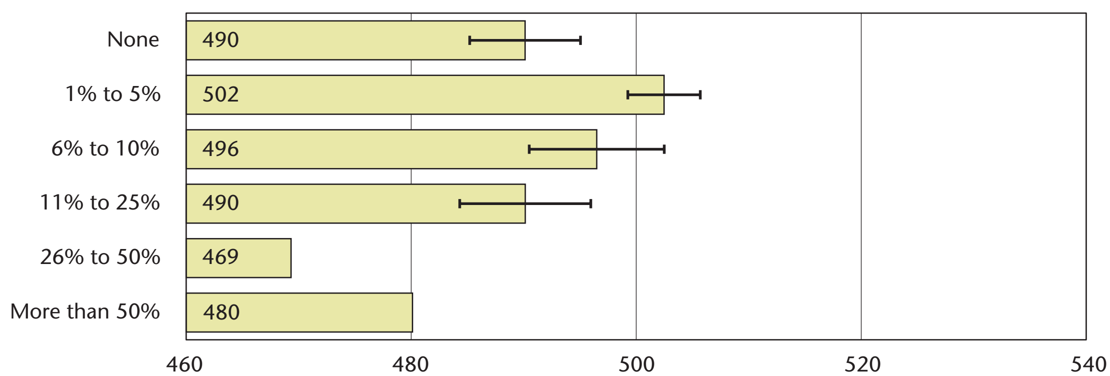
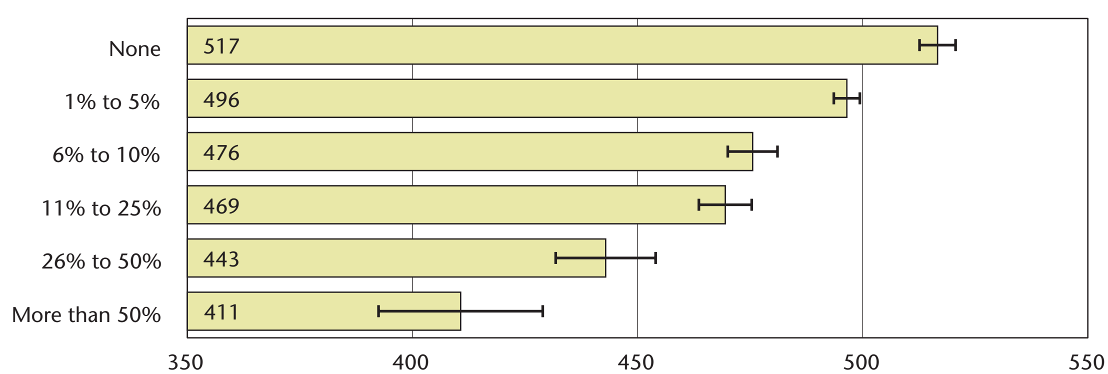


CHART 3-45 **Mean school mathematics scores by proportions of students of Aboriginal identity in publically funded schools**



School locations by community size

Chart 3-46 shows the percentages of schools in communities of various sizes by jurisdiction and language. This distribution reflects the overall proportion of the populations in various jurisdictions that are located in large urban versus small rural locations and is not directly linked to overall population size for the jurisdiction.

Mean school mathematics scores by community size are given in Chart 3-47. This shows a generally increasing performance trend as community size increases. However, there is no significant difference between communities in medium size compared to large cities.

Chart 3-46 **Percentage of schools by community size by jurisdiction and language**

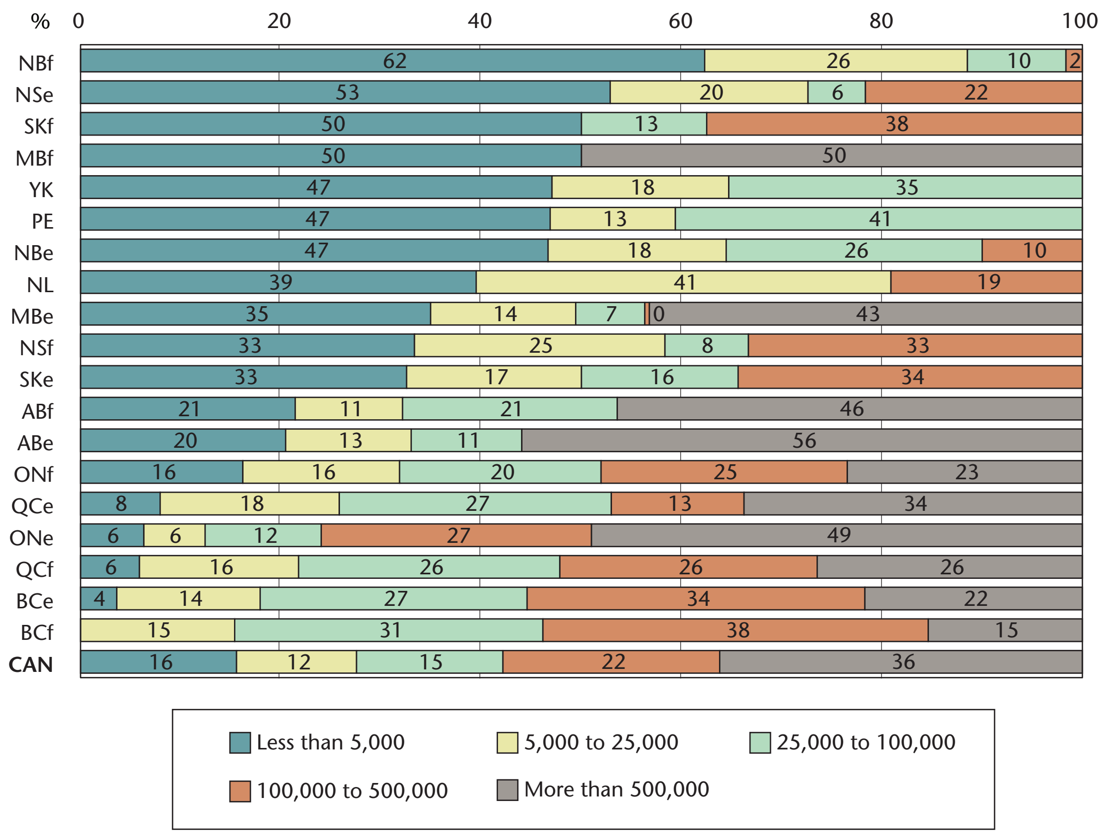
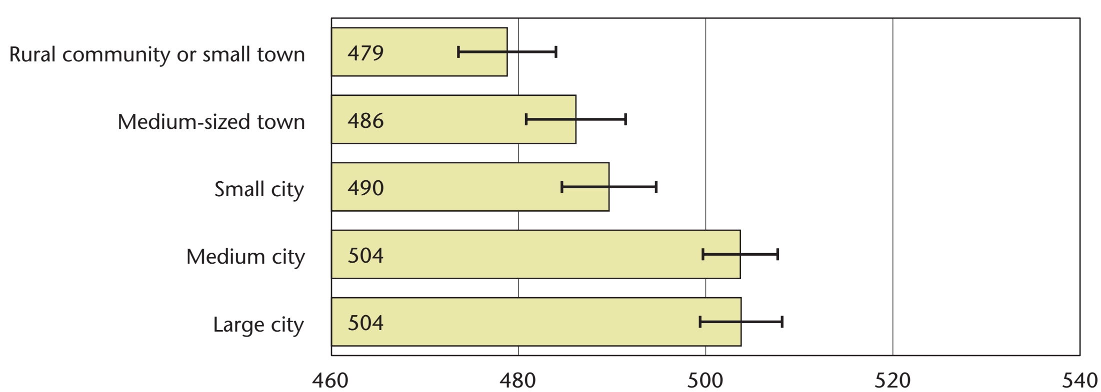


Chart 3-47 **Mean school mathematics scores by community size**



Statistical Note

Multiple Regression Analysis. Achievement is influenced by a large number of factors, which may act independently or in combination to affect the outcome. For example, results already presented indicate that both mother's education and the number of books in the home influence mathematics achievement. However, these two factors themselves are correlated. If taken together, one may be more prominent than the other or one may have no effect on achievement once the other is accounted for.

In survey research, the standard statistical technique for isolating effects is known as **multiple regression analysis or regression modelling**. This technique is based on an equation in which the outcome (or dependent variable) is seen as a linear combination of a series of factors (predictors or independent variables). The contribution of any one predictor to the outcome is represented by a **regression coefficient**, the value of which depends on the effect of the predictor itself and of the other variables in the model. The relative sizes of the regression coefficients in a particular model may be used to indicate the relative contributions of the factors of interest. Models which include or exclude a particular variable may also be used to identify the unique contribution of that variable while controlling for others.

Multilevel Modelling. The PCAP sampling model is a two-stage one, with schools sampled at a first stage and students within schools at a second. Multilevel modelling is a variation on regression analysis used in situations where the samples exhibit such a hierarchical structure. Models are developed at each level (i.e., the school level and student-within-school level), and the models are then combined to yield regression coefficients that represent effects at both the student level and school level. Most of the regression models used in this report are of this nature. For the most part, the results may be interpreted in the same way as for single-level models.

Interpreting Regression Coefficients. In general, a regression coefficient may be interpreted as representing the change in the outcome (in this case mathematics achievement) that would be expected from one unit change in a particular predictor (such as mother's education or amount of homework). **Simple regression** coefficients (sometimes called absolute effects) are those for the relationship between a single predictor and the outcome, without controlling for other variables. **Multiple regression** coefficients (sometimes called relative or unique effects) refer to the effects of a particular predictor while controlling for all other predictors in the equation.

The statistical significance of regression coefficients is determined from the confidence interval in the same way as described earlier. The specific reference point is a coefficient of zero, which would indicate that the factor has no correlation with the outcome variable. A coefficient may thus be said to be statistically greater than (or less than) zero if the error bar on the graph does not overlap the zero point. The absolute values of the coefficients for different variables cannot be compared directly in all cases because these depend on the scales used. We can say that one variable has a larger or smaller effect than another only if the two scales are the same. However, for any one predictor, the simple and multiple regression coefficients may be compared to determine the effect of controlling for other variables. This is the main comparison of interest for most of the models presented in this report.

To reduce the complexity of the models and graphs, not all variables examined in the earlier parts of each chapter are used in the multiple regression analysis. Generally, only those showing statistically significant effects or those judged to be of particular policy interest are included in the models.

In each chapter, the multiple regression coefficients represent the effects of each variable, controlling for all other variables for the specific model stage reported. For example, in Chapter 3, the multiple regression coefficient for gender represents the gender effect after controlling for all other student demographic variables. The final chapter presents a cumulative or "full" model, in which each multiple regression coefficient represents the effect for a specific variable, controlling for all other variables entered at previous stages.

Student variables

Chart 3-48 shows the simple and multiple regression effects on mathematics scores of student demographic variables. For dichotomous variables, the regression coefficient may be interpreted as the average difference in mathematics score between those possessing the characteristic and those not. For example, being male conveys a statistically significant 5.0 point advantage in mathematics over being female when the gender variable is taken alone. This advantage changes slightly to 7.3 points, when all other variables in the model are controlled. This is also statistically significantly greater than zero. However, the effect of controlling for the other demographic variables is not statistically significant, as evidenced by the overlap of the error bars for the two coefficients.

For variables with more than two values, the coefficient represents the effect of a change in mathematics scores of one point on the scale of the independent variable. For example, an increase of one unit on the "books in the home" scale conveys an advantage of 16.4 points in mathematics achievement when "books in the home" is taken alone. This does not change significantly when other student demographic variables are controlled.

At the student level, being male, using French or mainly French, higher educational aspirations, higher mother's education, and more books in the home are positively related to achievement. Being born in Canada and, using English or a language other than English or French, are negatively related to achievement. In all cases, these are statistically significant for both the simple and multiple regression models, indicating that these variables exert independent effects, which are not affected much by controlling for other variables.

School variables

Chart 3-49 shows the effects of school-level variables. Larger schools, private schools, or schools in larger communities have positive effects on mathematics achievement, while schools with greater student diversity, as inferred by a higher percentage of second language learners, have lower achievement. Again, the simple and multiple regression effects are similar, indicating that the effects of school demographic variables are largely independent of each other.

Teacher variables

The chart for the teacher variables is not shown because none of the teacher characteristics show statistically significant effects in the model.

Chart 3-48 **Regression coefficients for student demographic variables**[12]

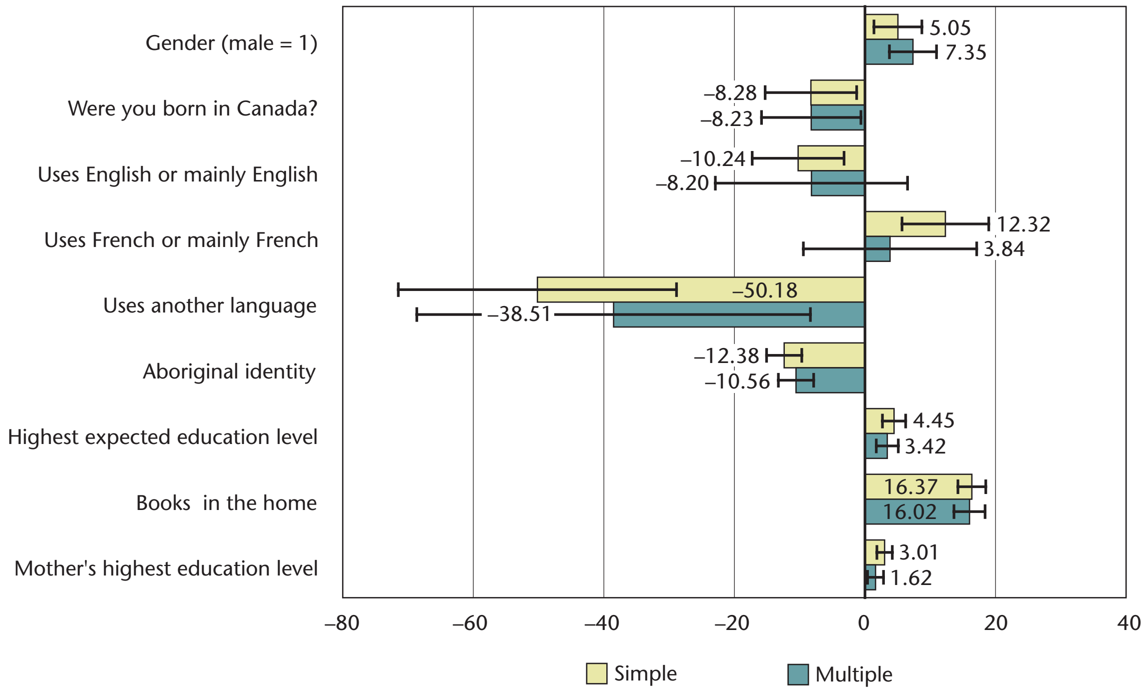
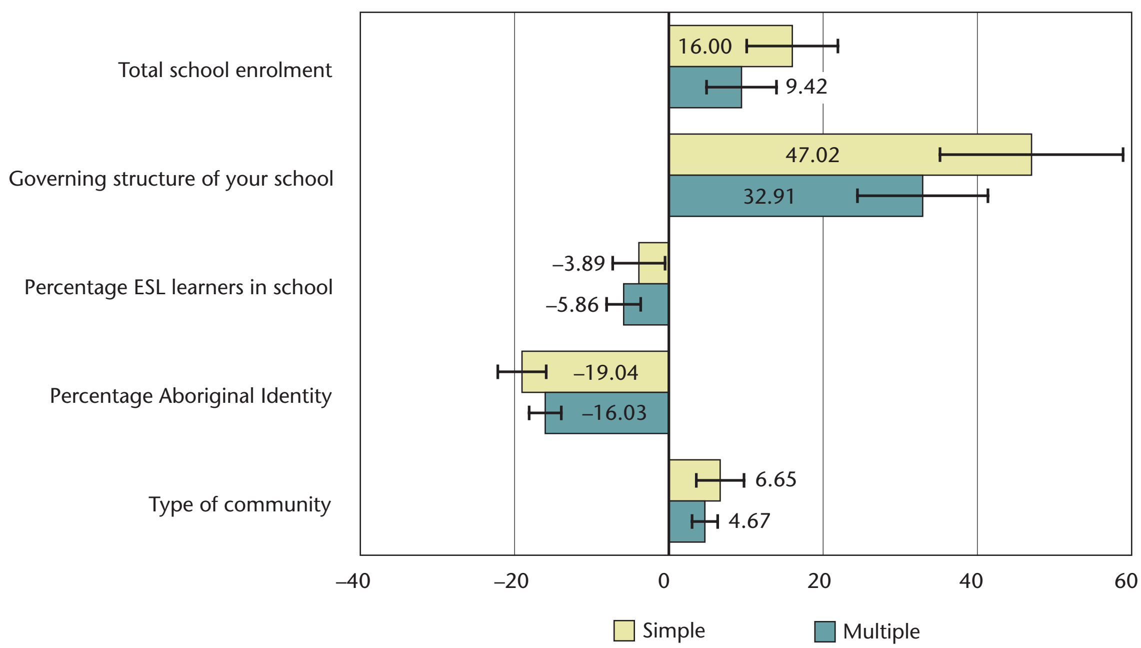


Chart 3-49 **Regression coefficients for school demographic variables**



[12] It is important to reiterate that the coefficients reported in the chart are not directly comparable across variables because the variables are on different scales. Effects for dichotomous variables are comparable because each represents simply values of zero or one. In other cases, the size of the effect depends on the number of categories on the scale. For any one variable, effects are comparable across the simple and multiple regression models.

4 STUDENT ATTITUDES

Statistical Note

Factor Analysis and Derived Variables

In order to reduce the complexity of the analysis and to obtain more stable measures of attitudes and behaviours, some groups of questions were subjected to factor analysis. This technique is designed to determine if item responses cluster together in some psychologically meaningful way. If meaningful groupings can be found, factor analysis permits the construction of a smaller number of factors or derived variables. For example, applying factor analysis to the student attitude questions yielded a set of seven derived variables, reduced from 30 individual questionnaire items. This illustrates the efficiency of this technique.

A "factor score" for each student on each derived variable was derived from the factor analysis, in much the same way as a scaled mathematics score was derived from analysis of the mathematics test items. Factor scores are typically computed in standard score form, with a mean of zero and a standard deviation of one. For convenience in presentation, and to avoid negative values on charts, the scores were transformed to a mean of 50 and a standard deviation of 10 for Canada as a whole. This is analogous to the transformation of mathematics scores to a mean of 500 and a standard deviation of 100. However, the scale is deliberately different to avoid confusion of factor scores with achievement scores. Mean factor scores for groups such as jurisdictions should be examined in relation to the Canada mean of 50 and standard deviation of 10. For example, a mean score of 52 for a group implies that the group is 0.20 standard deviation units above the mean for that factor. It is particularly important to stress that factor scores should not be interpreted as percentages.

It is noted that the names of derived variables are somewhat arbitrary but are intended to capture an underlying idea represented by the items that load heavily on a specific factor. Sometimes this is conveyed by a name similar to that of an item and in other cases the underlying idea is more generic. Included throughout the report are tables that identify questionnaire items with the corresponding derived variables. These are intended to convey a sense of how the factors have been labelled.

A number of questions to students were designed to obtain data on their attitudes toward school and toward mathematics. Questions were also asked about student attributions of success and failure, specifically on whether responsibility for success or failure is attributed to their own efforts (internal) or to others (external). This chapter examines the impact of student attitudes on mathematics scores and further develops the multiple regression models to account for attitudes.

Attitudes toward school

Five questions about how well students like school were included on the student questionnaire, using a conventional four-point scale from "strongly disagree" to "strongly agree". Responses to all of these questions were quite positive, with only small percentages in the "strongly disagree" and "disagree" categories.

The factor analysis yielded two factors from these five questions. The first three resulted in a factor labelled "liking for school" and the last two a factor labelled "Sense of belonging to school" as indicated in Table 4-1.

TABLE 4-1 **Questionnaire items for attitude toward school factors**

Factor	Items
Liking for school	• I like school • My teachers treat me fairly • My teachers care about me
Sense of belonging to school	• At school, I make friends easily • At school, I feel that I belong

Charts 4-1 and 4-2 give mean factor scores on each of these factors by jurisdiction and language. In Chart 4-1, students in the top five populations on the graph (Ontario English to Manitoba English) show greater liking for school than the Canadian average, while the lowest seven (Newfoundland and Labrador to Quebec French) are lower than the Canadian average. Quebec French is actually significantly lower than any other population. In Chart 4-2, the top four populations (Ontario French to Ontario English) are higher than the Canadian average, while the bottom nine (British Columbia English to Yukon) are lower than the Canadian average. It is noted that, although statistically significant, these differences are not particularly large, amounting to 0.20 standard deviation units above and below the mean.

CHART 4-1 **Mean factor scores for "liking for school" by jurisdiction and language**

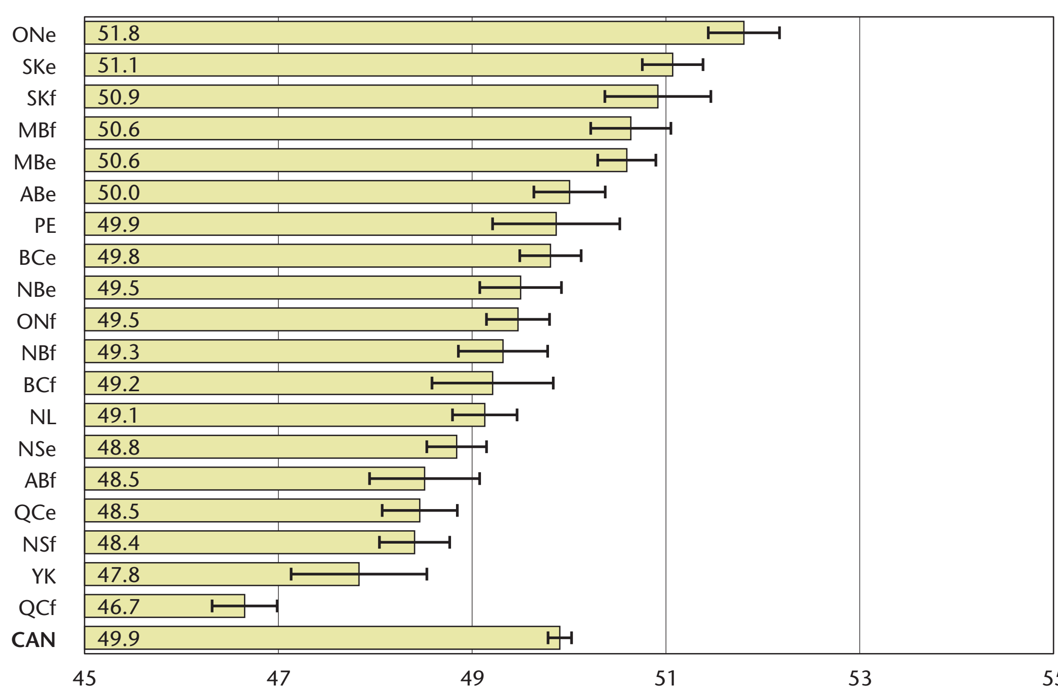
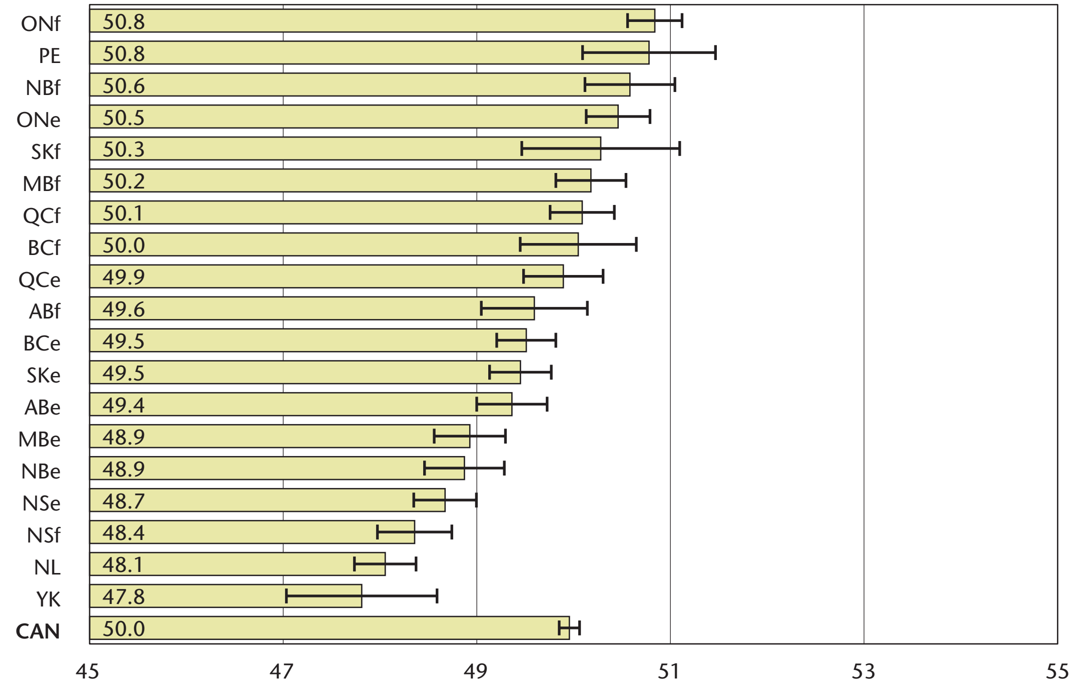


Chart 4-2 **Mean factor scores for "sense of belonging to school" by jurisdiction and language**

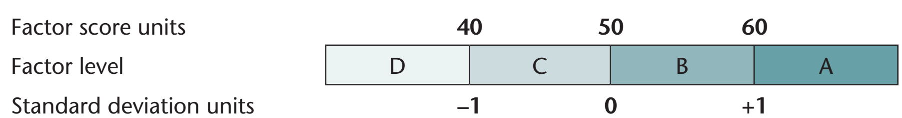


In order to examine the effects of these variables on mathematics achievement, students were divided into four "levels" based on their attitude factor scores, based on one standard deviation unit from the mean, as show in Table 4-2.[13] The factor levels are designated by letters to distinguish these from the achievement levels. Those in the highest level (A) have the most positive attitudes, and those in the lowest quartile (D) have the least positive attitudes.

Table 4-2 **Division of factor scores for analysis of mathematics achievement**

Factor Level	Standard Deviation Units	Factor Score Units (rounded)
D (lowest)	< –1 SD	Lower than 40
C (low)	–1 to 0 SD	40–50
B (high)	0 to +1 SD	50–60
A (highest)	> +1 SD	Higher than 60

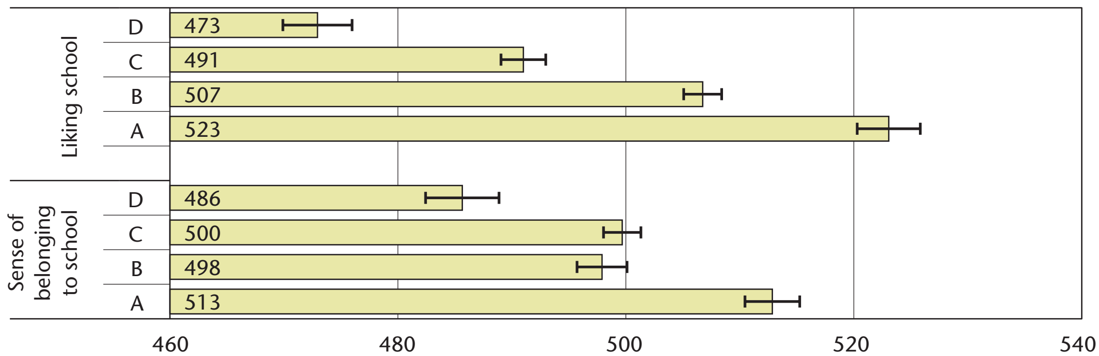
This can be represented visually as shown below.



[13] This is a change from the 2007 report, where the division was into "quintiles" or five equal-sized groups. This change was made to facilitate the reporting of model results in standard deviation units (10 points on the factor scale) rather than factor score units (one point on the factor score scale).

Chart 4-3 presents the effects on mathematics achievement of liking for school and sense of belonging to school. In this case, a general pattern of increased mathematics performance with more positive attitudes toward school is evident from both the levels and the means.

Chart 4-3 **Mean reading scores by attitudes toward school**



Attitudes toward mathematics

Students were asked to respond to 14 items on attitudes toward mathematics. Factor analysis of this item set yielded four factors, labelled as given in Table 4-3. The attitude that "mathematics is easy" was the strongest factor in this set. The second factor identifies students who like dealing with mathematics questions that involve a lot of reading. The third factor pertains to positive attitudes toward a variety of mathematics processes. Finally, the positive loadings on the fourth factor identify students who have negative views of mathematics.

Table 4-3 **Questionnaire items for attitude toward mathematics factors**

Factor	Items
Mathematics is easy	• Mathematics is an easy subject • I feel nervous when doing mathematics (–) • I understand most of the mathematics I am taught
Like mathematics questions with a lot of reading	• I like mathematics questions that do not require much reading (–) • I like mathematics questions that involve a lot of reading
Positive attitudes towards mathematics processes	• I like estimating • I like hands-on mathematics activities • I like doing mental mathematics • I like problem solving • I like doing paper-pencil calculations
Negative attitudes to mathematics	• Learning mathematics is a waste of time • The mathematics I learn now will be useful in the future (–) • I need to keep taking mathematics for the kind of job I want after I leave school (–) • Mathematics is boring

Note: (–) indicates items that showed negative loadings during the factor analysis. For example, in the model, a negative response to some items may yield positive factor scores. Negatively worded items, such as "I feel nervous when doing mathematics" have negative loadings on their respective factors, which is consistent with the positive label attached to the factor. This happens when the items correlate negatively with other items; however, items can have a positive loading on one factor and a negative loading on a different factor.

Mean scores by jurisdiction and language for these four factors are given in Charts 4-4 to 4-7. The following highlights may be noted:

- On the "mathematics is easy" scale, most of the francophone populations are significantly above the Canadian average. The exception is Quebec French, which is near the lowest on the scale.
- On the same scale, most other jurisdictions are significantly below the Canadian average.
- Differences on the "mathematics reading" scale are smaller than on the "easy" scale, with about half the populations at or near the Canadian average.
- On the "process" scale, Ontario English stands out as higher and Quebec French as lower than any other population.
- On the "negative" scale, higher numbers are interpreted as more negative attitudes. Students in Prince Edward Island and New Brunswick English stand out as significantly higher than most other populations. Four populations, including both Ontario populations, along with Nova Scotia French and New Brunswick French, have lower than average scores on the negative factor.

CHART 4-4 **Mean factor scores for "mathematics is easy" by jurisdiction and language**

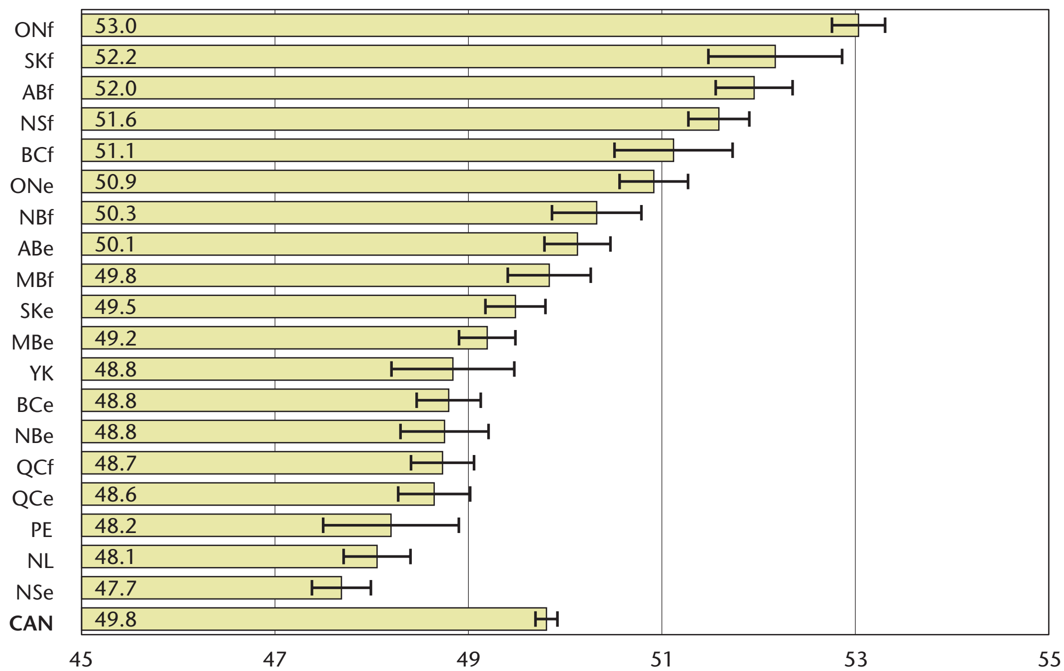


CHART 4-5 **Mean factor scores for "like mathematics questions with a lot of reading" by jurisdiction and language**

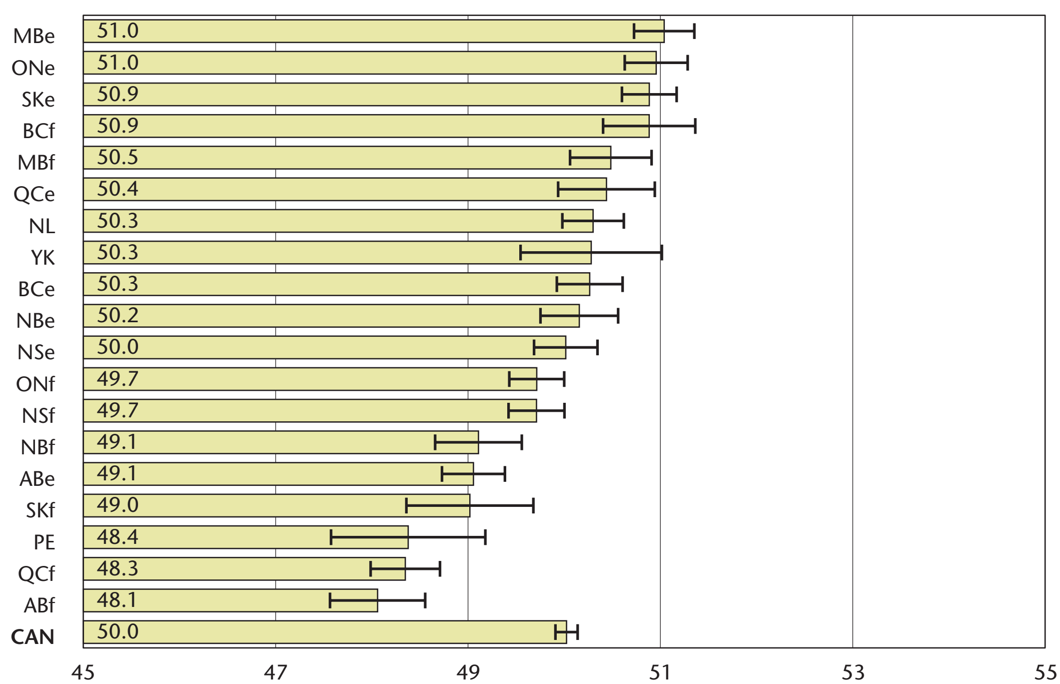


Chart 4-6 **Mean factor scores for "positive attitude towards mathematics processes" by jurisdiction and language**

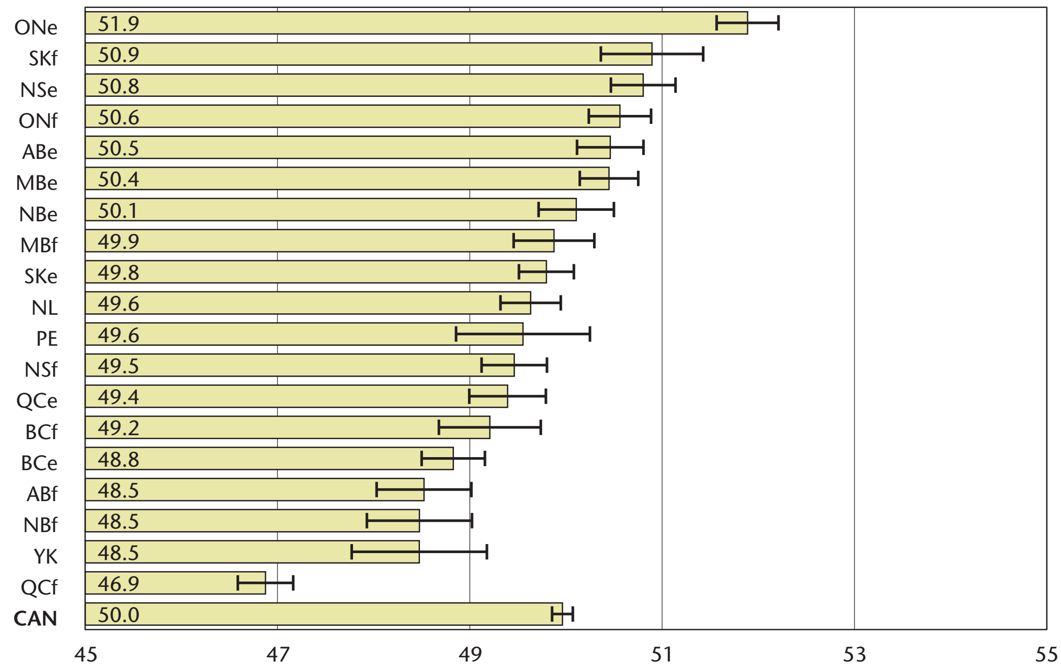
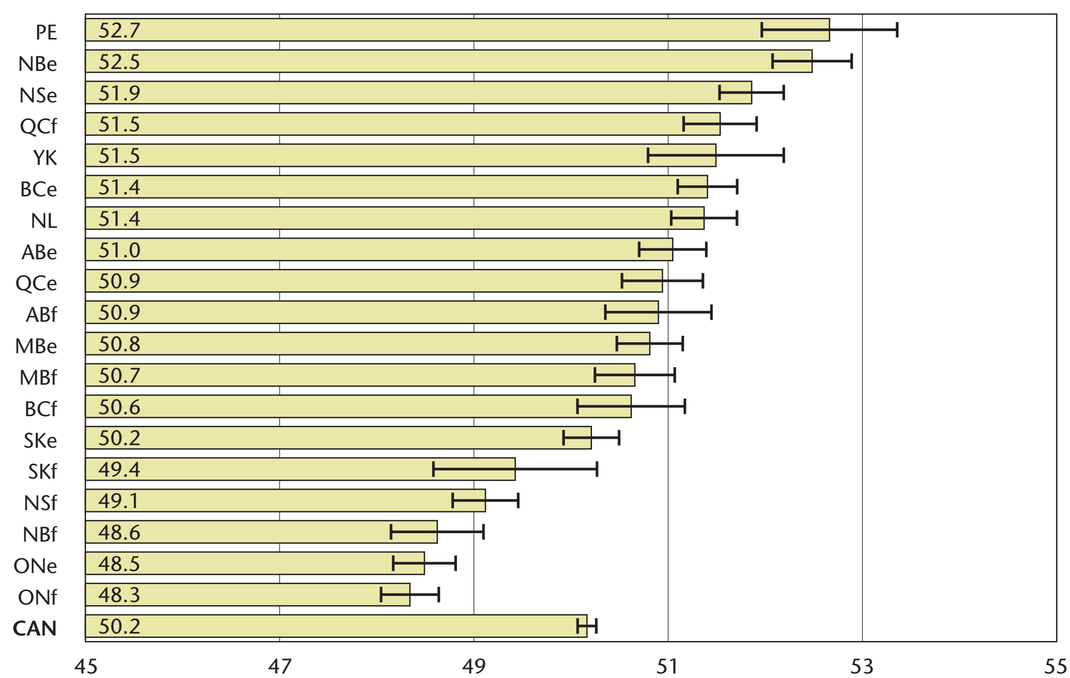


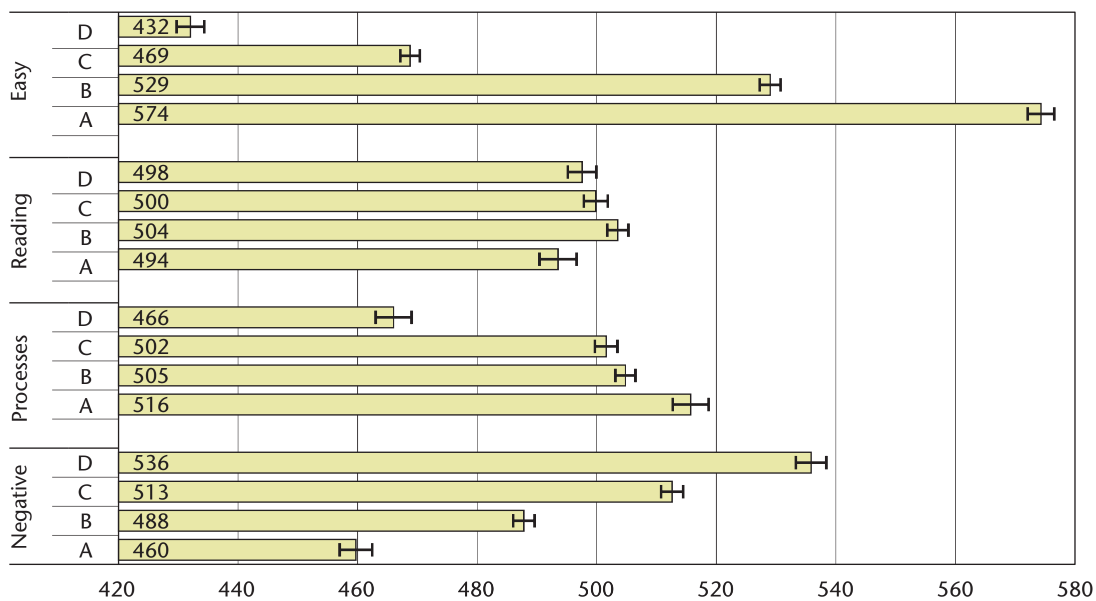
Chart 4-7 **Mean factor scores for "negative attitudes to mathematics" by jurisdiction and language**



Mathematics mean scores for the four levels on these four factors are shown in Chart 4-8. The patterns for "easy" and "negative" are quite strong, with those who find mathematics easier having higher scores and those with more negative attitudes having lower scores. Those with higher scores on the process factor also tend to have higher scores in mathematics. Differences on the "mathematics reading" factor are smaller, but those at the highest level on this factor do less well than those at other levels.

An appropriate point to reiterate is that these are simple regression relationships and that no causal direction can be inferred from these results. In some cases, such as gender or socioeconomic status, there is only one plausible causal direction, with these factors affecting achievement rather than the other way around. However, for attitudes, there is no easy way to determine from the results, or from more basic temporal or other patterns, whether more positive attitudes lead to higher achievement or higher achievement yields more positive attitudes. Nevertheless, since improved achievement, rather than improved attitudes may be considered as the more fundamental goal of schooling, it is perhaps more plausible to argue that schools should strive to improve attitudes as a possible route to improved achievement, rather than strive to improve achievement in order to engender better attitudes.

Chart 4-8 **Mean mathematics scores by attitudes toward mathematics**



Attributions of success and failure

The next set of student attitude items had to do with their attributions of success and failure in their mathematics school work. Factor analysis of these items yielded four factors, as described in Table 4-4. Two factors were derived for external attributions, one for failure and one for success. These are labelled "negative" and "positive" respectively. Although these may be seen as opposite attributes, this was not revealed by the factor analysis. A third factor, labelled "fatalism," is related to attributing success or failure to luck rather than ability. Finally, the fourth factor is labelled "ability" because the reference items are positive attribution of success to natural ability and negative attribution of success to tutoring. This pattern is slightly different from typical attribution patterns, for which more distinct "internal" versus "external" attributions of success and failure tend to be found.

Table 4-4 **Questionnaire items for attribution factors**

Factor	Items
Negative	• No encouragement from my parents/guardians • No encouragement from my friends • No help with homework outside of school • Poor teaching • Not working hard enough
Positive	• Encouragement from my friends • Encouragement from my parents/guardians • Working especially hard • Good teaching
Fatalism	• Bad luck (–) • Good luck • Not enough natural ability (–)
Ability	• Natural ability • Tutoring outside of school (–)

Note: (–) indicates items that showed negative loadings during the factor analysis.

Charts 4-9 to 4-12 give mean factor scores on these factors by population. The following are some highlights of these results:

- Students in francophone populations tend to have lower scores on the negative attribution scale than those in anglophone populations. However, the same pattern is not evident on the positive attribution scale.
- Most populations are above average on the fatalism scale. The notable exceptions are both Ontario populations and Saskatchewan French.
- These same three populations (both Ontario populations and Saskatchewan French) are also near the top of the scale for the ability factor. Ontario French stands out as higher than any other population on this scale.

Chart 4-9 **Mean factor scores for negative attributions by jurisdiction and language**

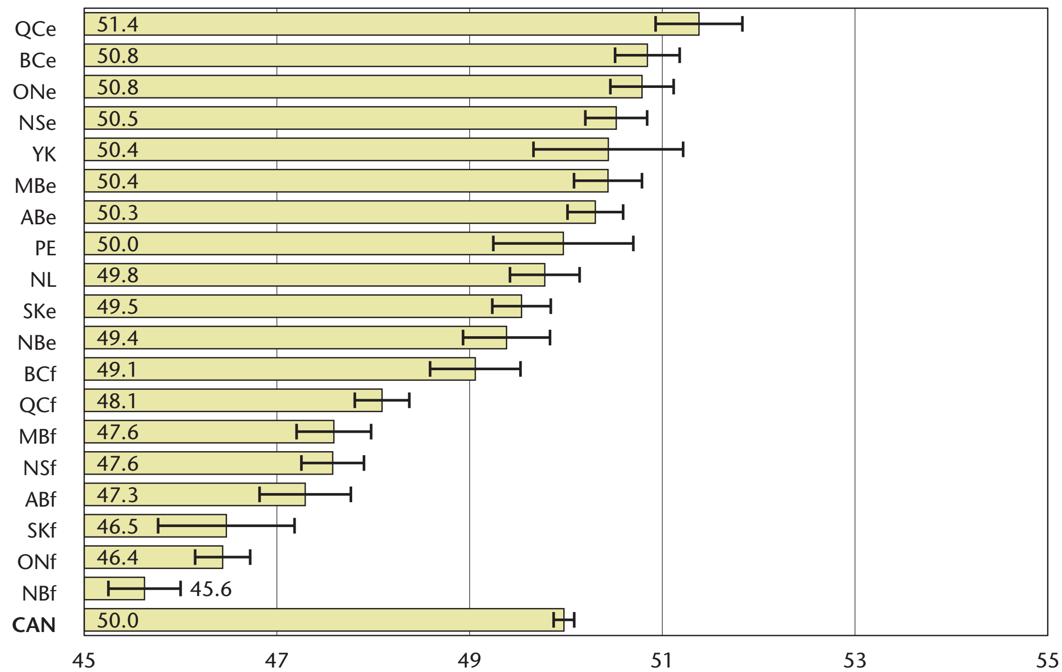


Chart 4-10 **Mean factor scores for positive attributions by jurisdiction and language**

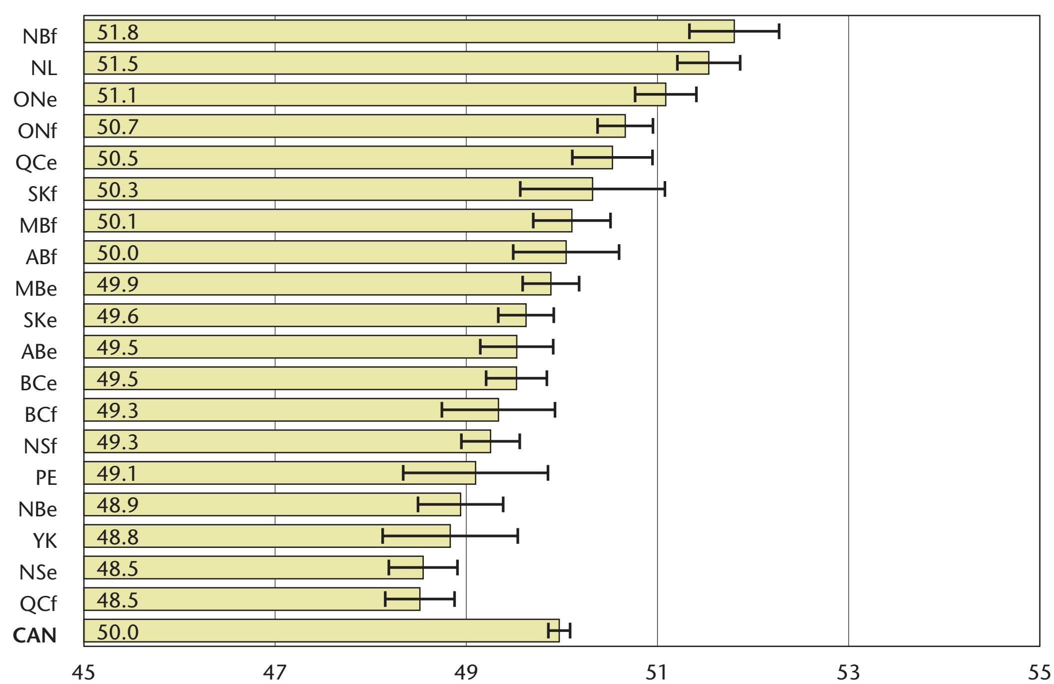


Chart 4-11 **Mean factor scores for fatalism by jurisdiction and language**

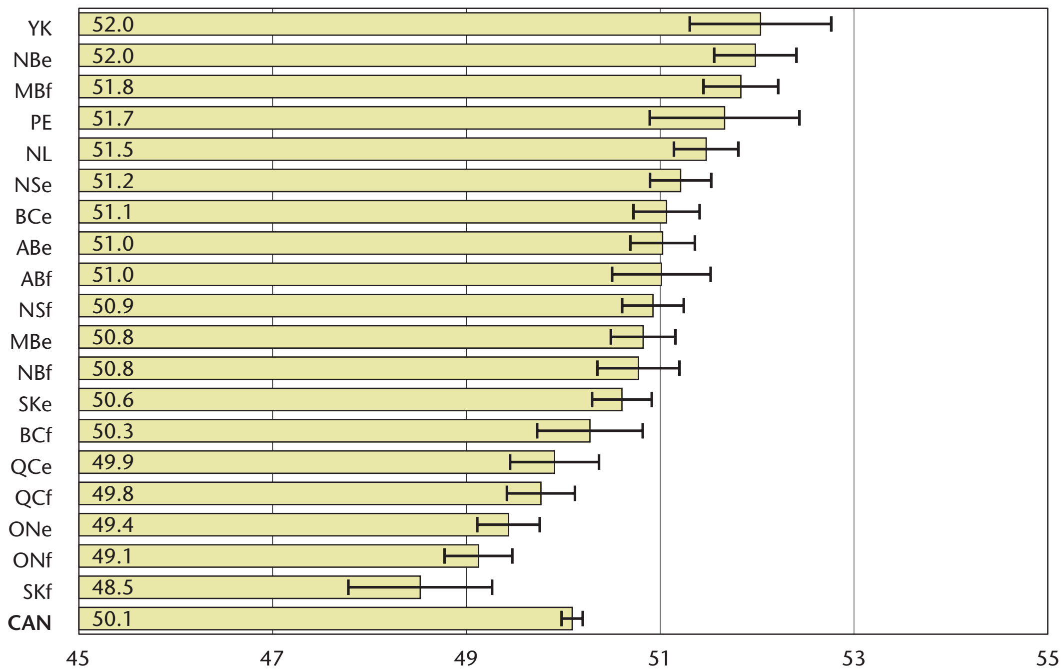


Chart 4-12 **Mean factor scores for attributions of success to ability by jurisdiction and language**

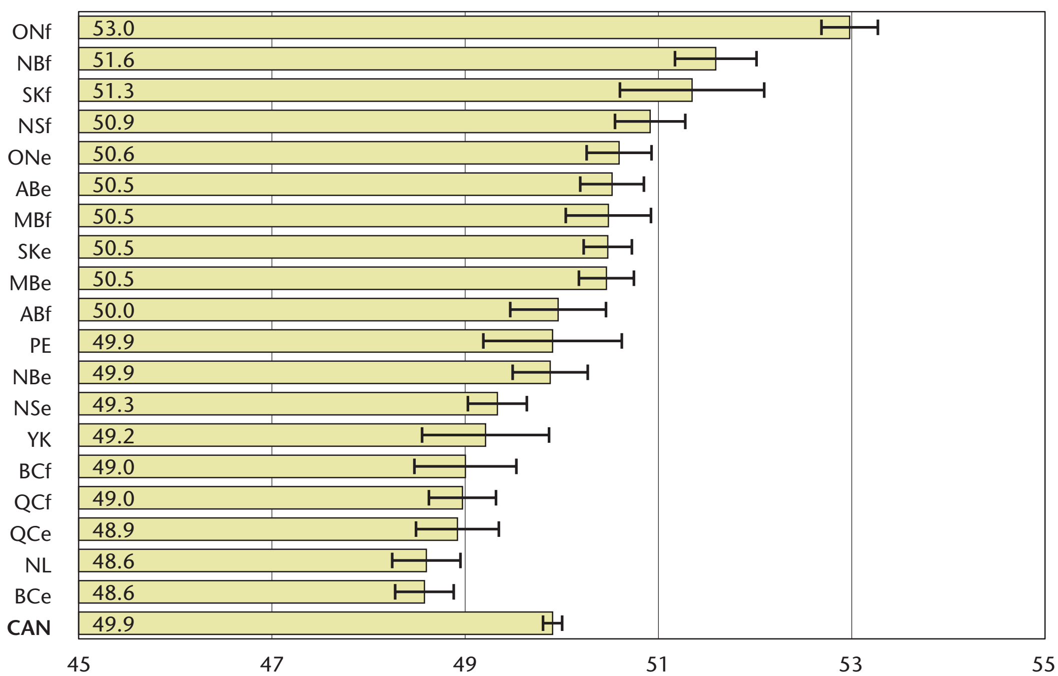
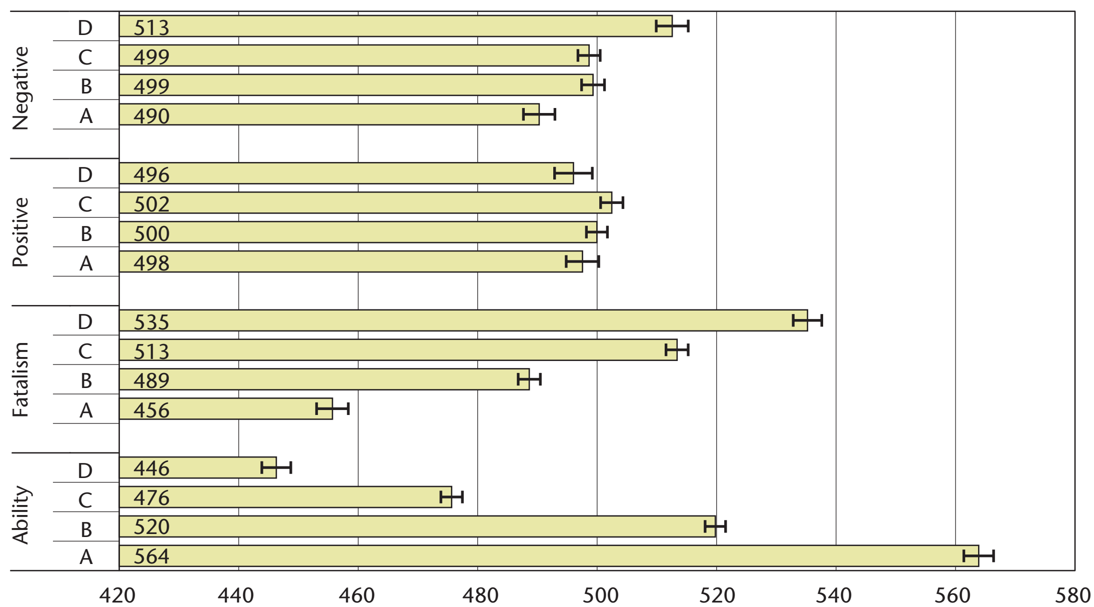


Chart 4-13 gives mean mathematics scores by levels of these factors. The patterns for fatalism and ability are quite clear, with higher levels of fatalism associated with lower mathematics scores and higher levels of ability associated with higher mathematics scores. The pattern for negative attributions shows lower scores associated with more negative attributions. There are no significant differences in the mean mathematics scores on the positive attribution factor.

Chart 4-13 **Mean mathematics scores by attribution factors**



Confidence in mathematics

A set of 10 questions on confidence in doing mathematics resolved into three factors as indicated in Table 4-5.

Table 4-5 **Questionnaire items for confidence factors**

Factors	Items
General confidence in mathematics	• Confidence with: – Mental math – Paper-pencil calculations – Problem solving – In general, how confident do you feel in mathematics? – Estimation – Reading to solve problems
Computing confidence	• Confidence in using computers in mathematics • Confidence in using calculators in mathematics
Decreased confidence over time	• When did you feel the most confident in mathematics? (–) • Since you started school, how would you say your level of confidence in mathematics has changed?
Note: (–) indicates items that showed negative loadings during the factor analysis.	

Charts 4-14 to 4-16 give the distribution of mean confidence factor scores by population. For general confidence, two high performing populations are at opposite ends of this scale, Ontario English at the top and Quebec French at the bottom, both being significantly different from all other populations. Alberta French, New Brunswick French, and Yukon are the only other populations below the Canadian average on this factor, while most others are close to the Canadian average. On the computing confidence scale, only Ontario English and Quebec English are above the Canadian average, while about half the populations are below. Finally, on the decreased confidence scale, most of the francophone populations are below the Canadian average.

Chart 4-14 **Mean factor scores for general confidence in mathematics by jurisdiction and language**

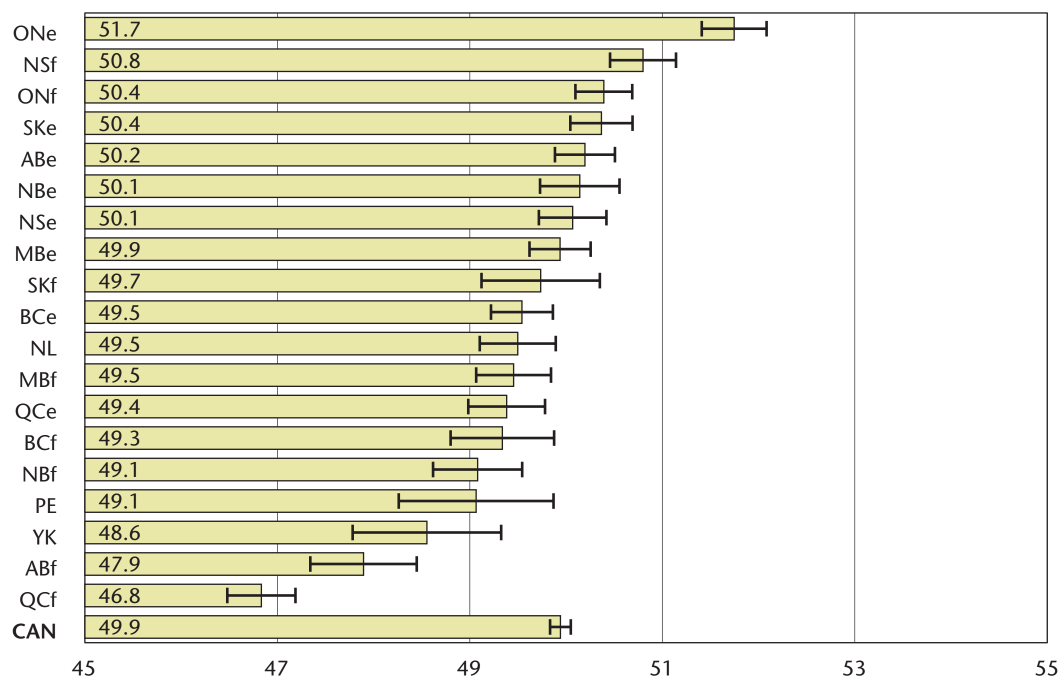


Chart 4-15 **Mean factor scores for computing confidence by jurisdiction and language**

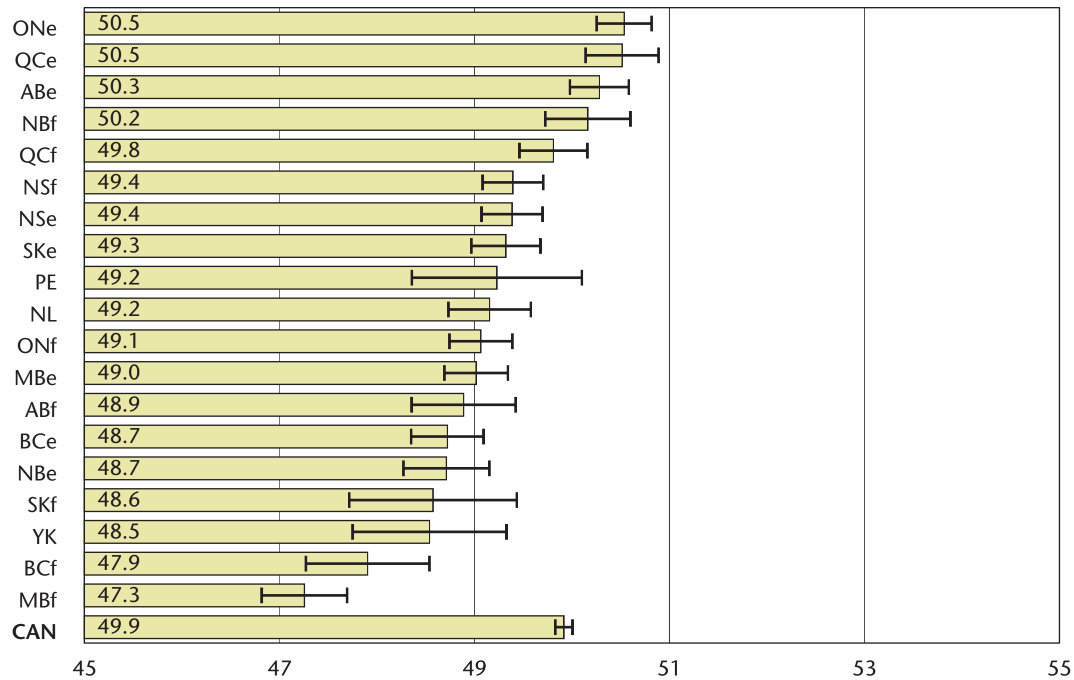


Chart 4-16 **Mean factor scores for decreased confidence over time in mathematics by jurisdiction and language**

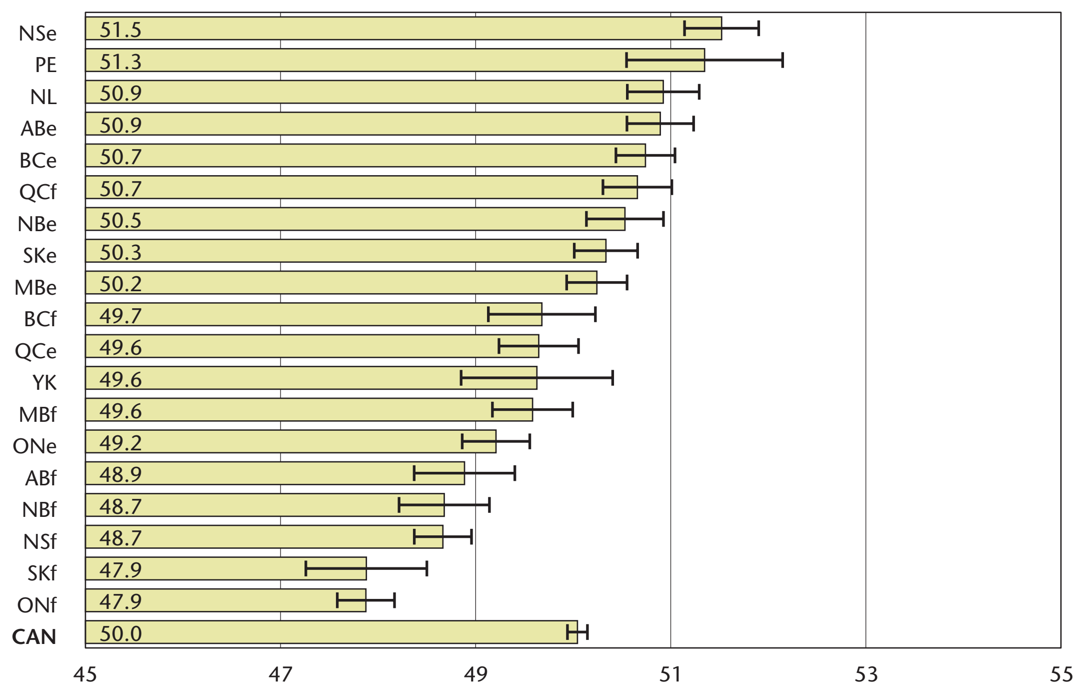
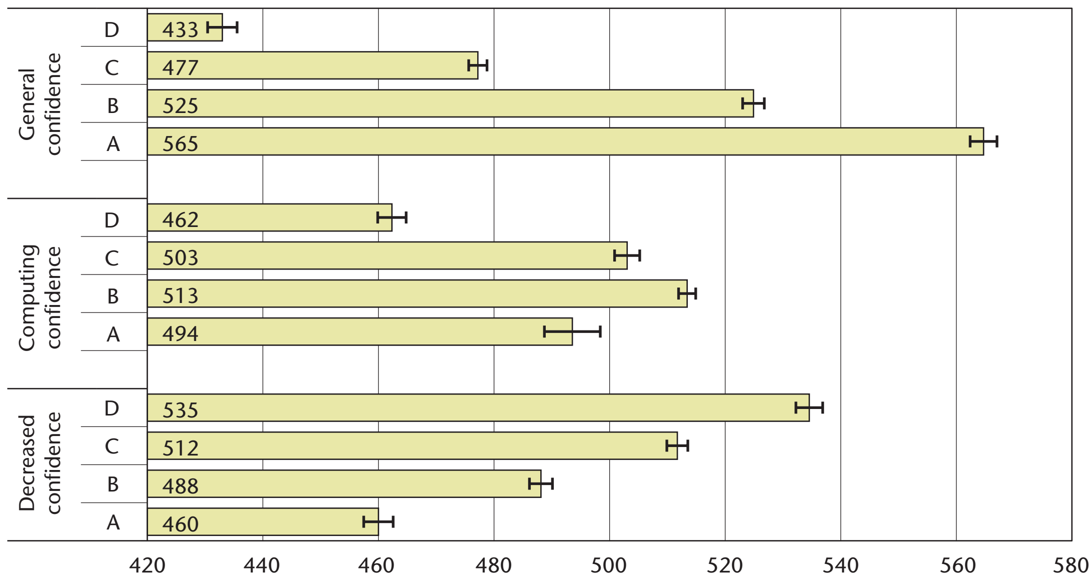


Chart 4-17 gives mean mathematics scores for the four levels of these factors. This shows a strong pattern of higher mathematics scores with increased general confidence and lower mathematics scores with decreased confidence. Higher computing confidence shows a non-linear pattern, with increasing mean mathematics scores up to the third level, and a significant decrease at Level A compared to levels B and C.

Chart 4-17 **Mean mathematics scores by confidence factors**



Multiple regression effects

The effects of attitudes on mathematics achievement were modelled by examining simple and multiple regression coefficients based on a two-level model as in the previous chapter. Here the simple regression coefficients represent the change in mathematics score of a one standard deviation (10 points on the factor score scale) change in one of the attitude variables. The multiple regression coefficients represent the change in mathematics score of a one standard deviation change in a particular attitude variable, controlling for all of the other attitude variables.[14] In this case the coefficients are directly comparable across variables as well as across the simple and multiple regression models because all of these variables are on the same factor score scale.

In general, the simple regression effects show the same pattern as revealed by the previous comparison of mean scores by standard deviation units. All of the simple regression effects, except for liking for mathematics problems with a lot of reading and positive attributions, are statistically significant, with negative attitudes, fatalism, and decreased confidence over time being negative. However, differences in the relative sizes of the various effects are more apparent here. Mathematics is easy, general confidence in mathematics and attribution of success to ability have by far the largest positive effects. Negative attitudes to mathematics, decreased confidence over time, and fatalism have relatively large negative effects.

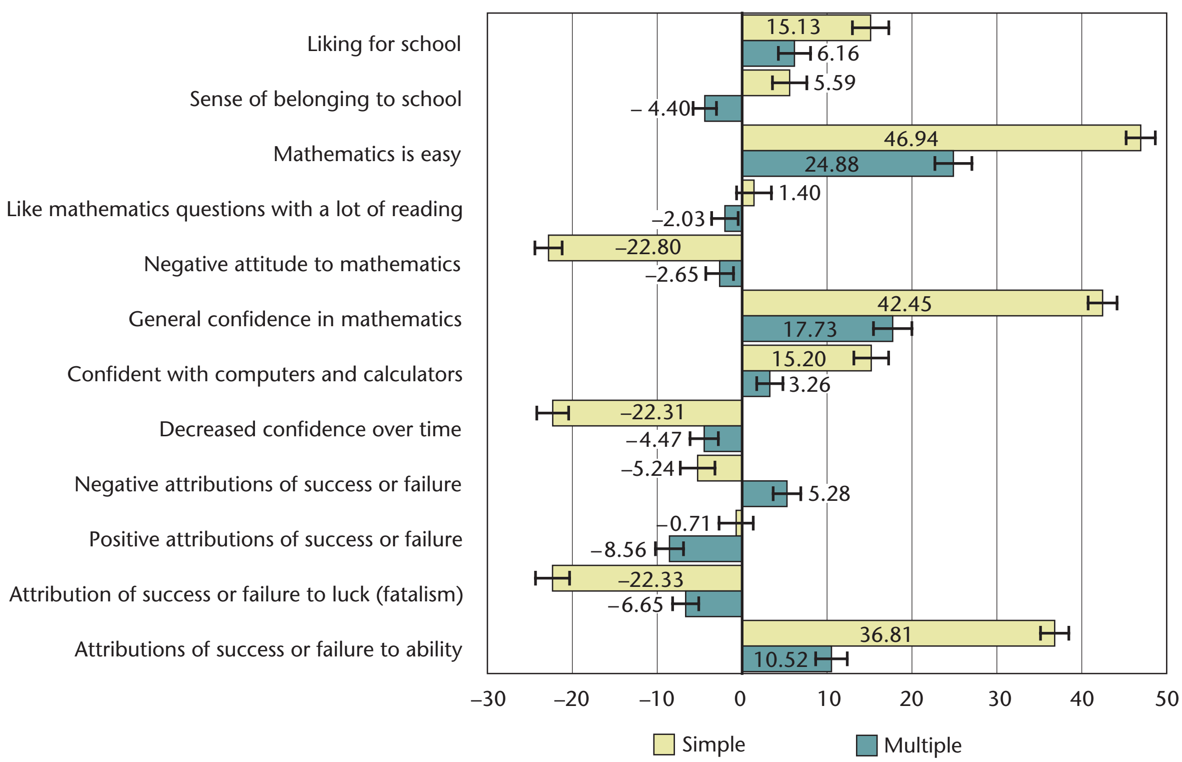
[14] This represents two changes from the approach taken in the PCAP 2007 Contextual Report. First, the change in achievement is now for one standard deviation unit (10 score points) rather than one score point of the factors. Second, in 2007, at each stage of the model, the variables for all earlier stages were controlled. Thus, in 2007, the models for attitude variables also controlled for demographic variables. In this report, the variables for each stage are examined as a group before reporting, in the last chapter, a full model with all variables controlled.

Almost all of the effects are significantly attenuated in the multiple regression model relative to those in the simple regression model, though most remain statistically greater than zero. Indeed, for two of the variables, belonging to school and negative attributions, the direction of the effects are reversed in the multiple regression model. This indicates that the effect for any one attitude variable is related in some way to the effects of the other variables in the model.

Since these effects sizes are directly comparable, it can be said that the largest positive effects are for the factors labelled mathematics is easy, attribution of success or failure to ability, and general confidence in mathematics. The largest negative effects are for negative attitudes toward mathematics, fatalism, and decreased confidence over time. All of these effects are significantly attenuated in the multiple regression model, which indicates that these effects are offset to some degree by other more positive attitudes.

These effects may be interpreted directly in terms of the impact on mathematics achievement for one standard deviation change (10 score points) in the attitude variable. For example, a 10-point change on the "mathematics is easy" scale contributes to about a 47-point change in mathematics scores when taken alone and a 25-point change in achievement even when all other attitude variables are controlled. Similarly, a 10-point decrease in confidence over time contributes to a 22-point decrease in mathematics achievement, which reduces to a 4-point decrease once other attitudes are controlled. In general, the combination of positive attitudes appears to have a greater positive effect on achievement than the negative effects of the combination of negative attitudes.

CHART 4-18 **Regression coefficients for student attitude variables**



5 STUDENT MATHEMATICS BEHAVIOURS AND STRATEGIES

This chapter examines the impact of mathematics-related activities and strategies on mathematics achievement. Questions from several scales on the student questionnaire were factor analyzed, with meaningful factor patterns being found in each case. Mathematics scores were examined in relation to these factors, and these scores were modelled by including these factors in the two-level regression equations, controlling for all of the other variables in each set.

Strategies on encountering difficult mathematics problems

A set of nine questions was designed to capture how students react when they encounter difficult mathematics problems. Factor analysis of these items yielded three strategies, as identified in Table 5-1.

Mean factor scores by population for these factors are given in Charts 5-1 to 5-3. For the persistence factor, only one population, Ontario English, is above the Canadian average, though several are below. Nova Scotia English stands out as being lower than any other population. For the on-line help factor, Ontario English, Quebec English, and British Columbia English students are above the Canadian average, while most others are below. In this case, Saskatchewan French is lower than any other population, despite its large margin of error. Finally, on the seeking help from others factor, Newfoundland and Labrador, Saskatchewan French, and British Columbia English are above the national average and also above most other populations. Most of the francophone populations are below the Canadian average on this factor.

Table 5-1 **Strategies when encountering difficult mathematics problems questionnaire items and factors**

Factors	When I encounter difficulty with mathematics I…
Persistence	• try several ways until I find one that works • give up (–) • look at the answer, if it is available, to see if that gives me a clue about how to do the work • look for examples in textbooks or notes
On-line help	• do a search on the Web • check a computer help site (tutoring site)
Help from others	• look for help from a classmate/friend • look for help from my teacher • look for help at home
Note: (–) indicates items that showed negative loadings during the factor analysis.	

Chart 5-1 **Mean factor scores for persistence by jurisdiction and language**

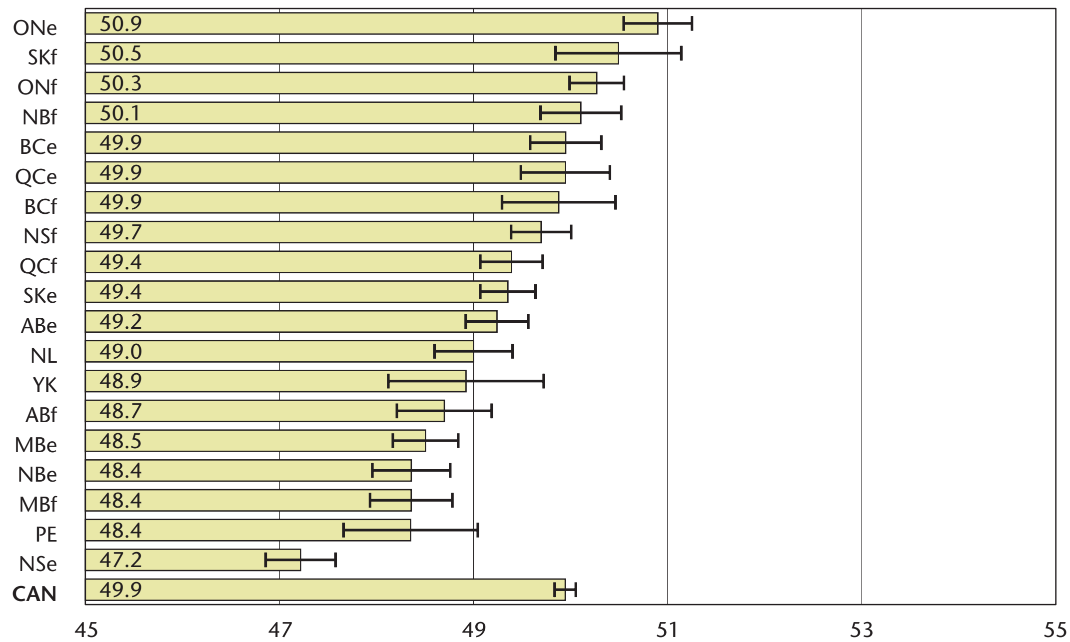


Chart 5-2 **Mean factor scores for on-line help in mathematics by jurisdiction and language**

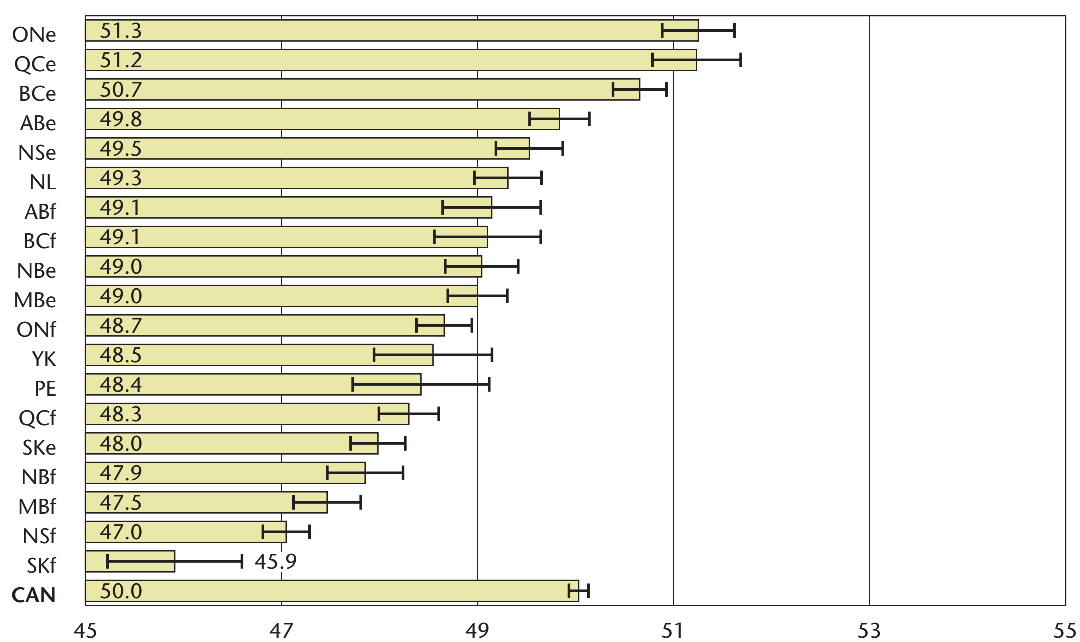
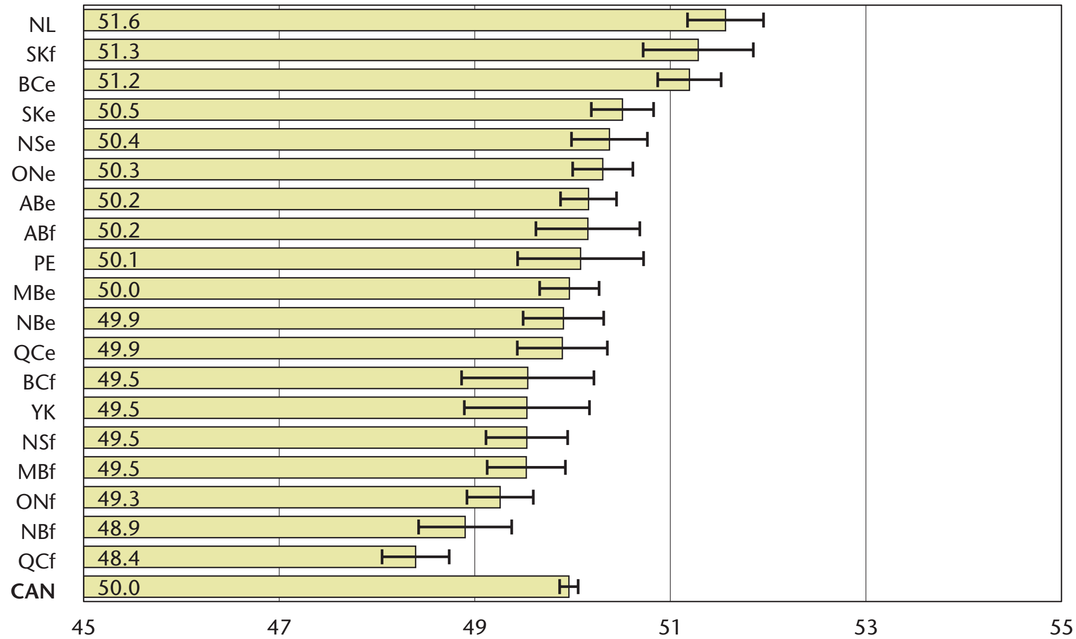
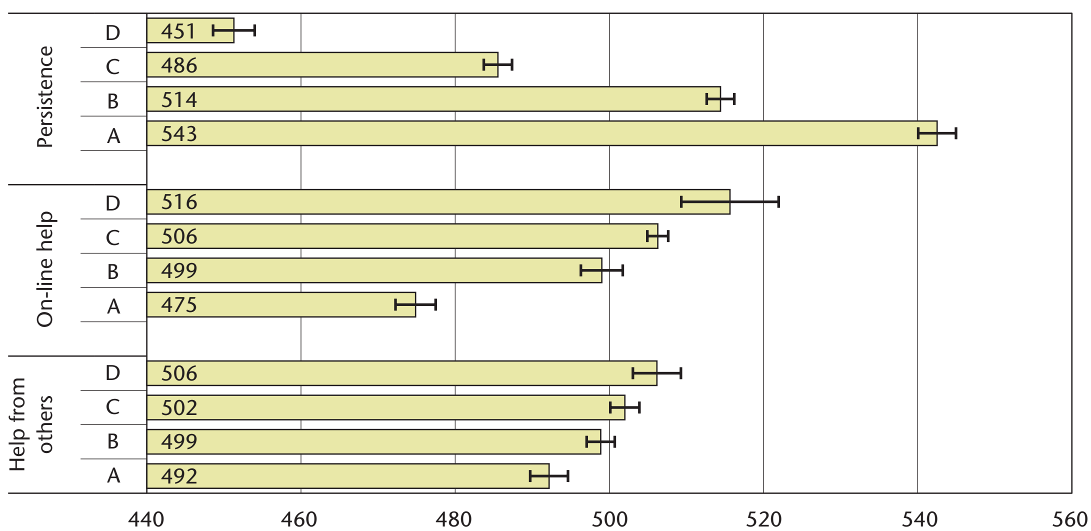


Chart 5-3 **Mean factor scores for seeking help from others in mathematics by jurisdiction and language**



For purposes of examining their effects on mathematics achievement, the factor scores were again divided into four categories based on standard deviation units, as described in Chapter 4. Chart 5-4 shows the mean mathematics scores by category (D is the lowest and A the highest) for these three variables. A clear pattern emerges here, with more persistent students having higher mathematics scores and those who more often seek either on-line help or help from others having lower scores.

Chart 5-4 **Mean mathematics scores by strategies for dealing with difficult mathematics problems**



Time on out-of-school activities

A 10-item question set was designed to gather data on activities outside of school that might relate to school achievement. These items were on a 6-point time scale, from no time to more than 6 hours per week. This set of items yielded four factors, as shown in Table 5-2.

Table 5-2 **Out-of-school activities questionnaire items and factors**

Factors	Items
Outside help	• Working with a mathematics tutor • Getting extra help at school, outside of regular school hours • Using a computer for school purposes (e.g., research, writing)
Entertainment using technology	• Playing computer, video, or other electronic games • Watching television or movies • Using a computer for personal reasons (e.g., Internet, e-mail) • Playing mathematics-related games
Sports/outside lessons	• Doing sports or other school and community activities • Taking other lessons (e.g., music, swimming) • Playing mathematics-related games
Personal communications	• Using the telephone or texting • Using a computer for personal reasons (e.g., Internet, e-mail) • Playing mathematics-related games (–)
Note: (–) indicates items that showed negative loadings during the factor analysis.	

Mean factor scores for these variables by population are given in Chart 5-5 to 5-8. For the outside help factor, Quebec English and Newfoundland and Labrador stand out as higher than any other jurisdiction. Ontario English and British Columbia English are lower than the first two but higher than the Canada mean. Most other populations are below the Canada mean. On the entertainment using technology factor, Yukon students are higher than any others, with most others being above the Canada mean and only Ontario English being below. For the sports/outside lessons factor, several populations from Newfoundland and Labrador to Quebec French cluster at the top, as well as Manitoba English, with means significantly higher than the Canada mean. Most of the francophone populations, along with Ontario English, cluster below the Canada mean. The opposite seems to be true for the personal communication factor, where four of the francophone populations, along with Manitoba English, British Columbia English, and Alberta English, are in the top cluster, significantly higher than the Canada mean, while Newfoundland and Labrador, Nova Scotia English, and New Brunswick English are in the bottom group.[15]

[15] The factor analysis technique used allows for some correlation between the factors for a particular set of items. On a time scale, such as the one used for these items, negative correlations between the factors might be expected because there is a limited amount of total time available for these activities. Time spent on one kind of activity is thus likely to be at the expense of others. Thus populations near the top on one factor are likely to be near the bottom on others.

Chart 5-5 Mean factor scores for outside help with school work by jurisdiction and language

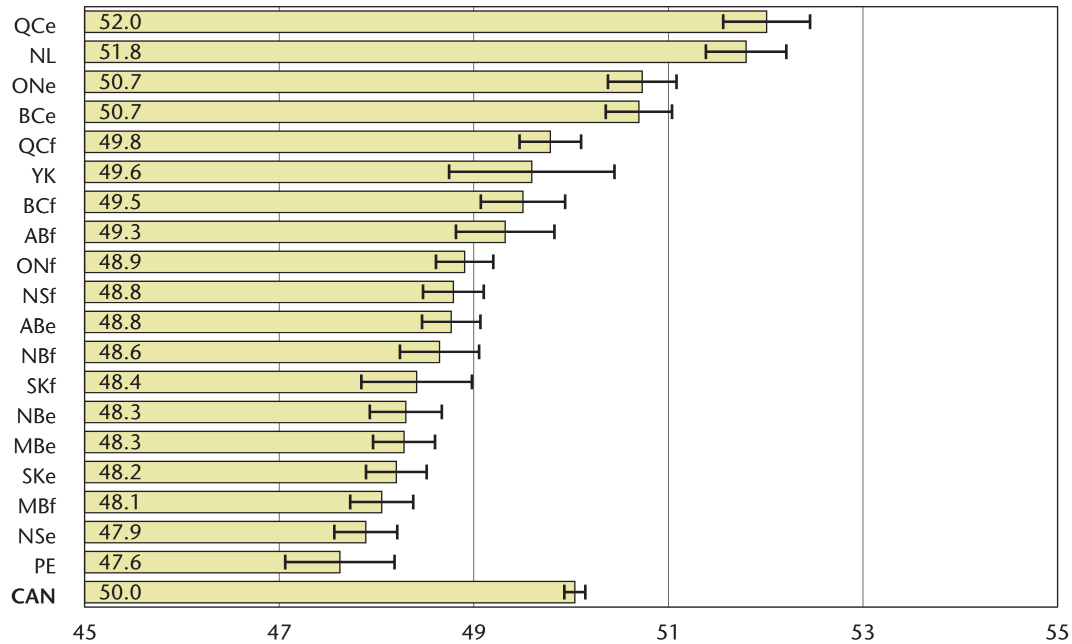


Chart 5-6 Mean factor scores for entertainment using technology by jurisdiction and language

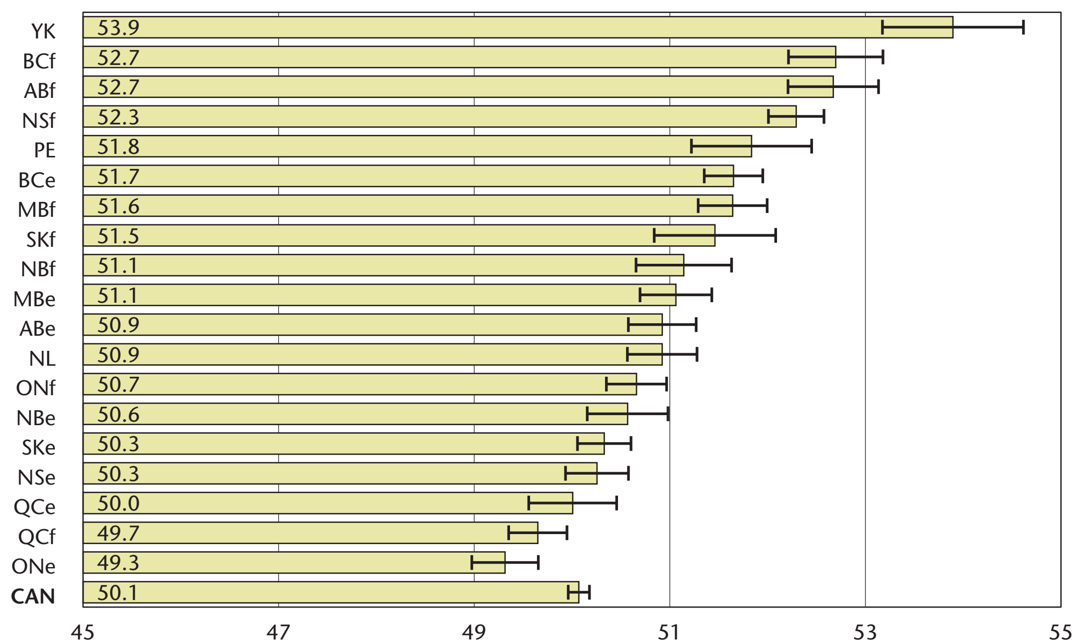


CHART 5-7 **Mean factor scores for sports/outside lessons by jurisdiction and language**

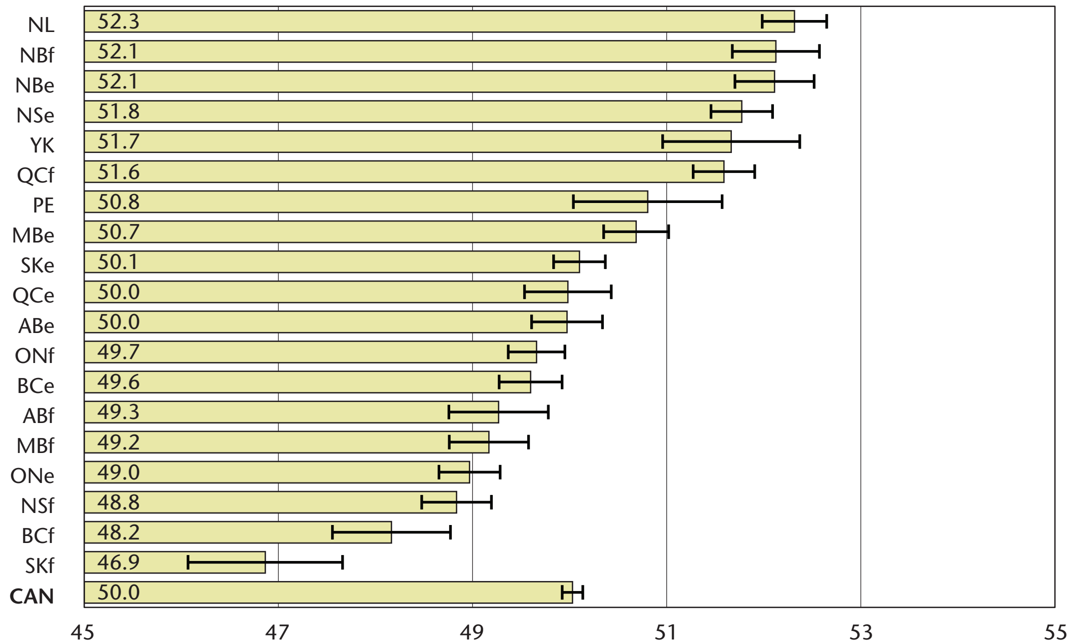


CHART 5-8 **Mean factor scores for personal communications by jurisdiction and language**

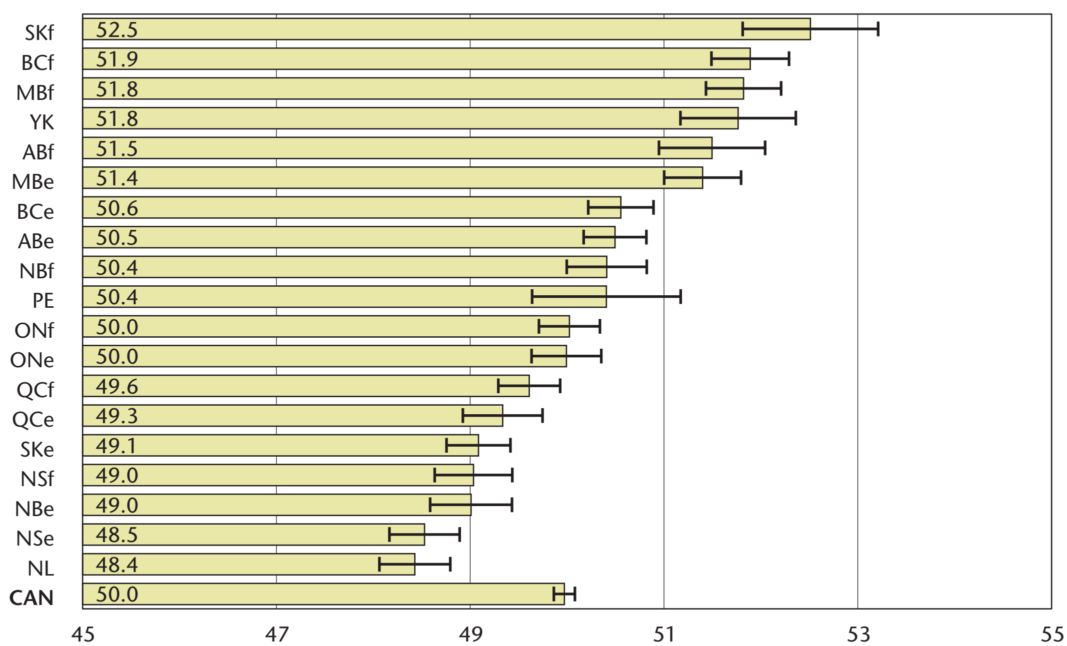
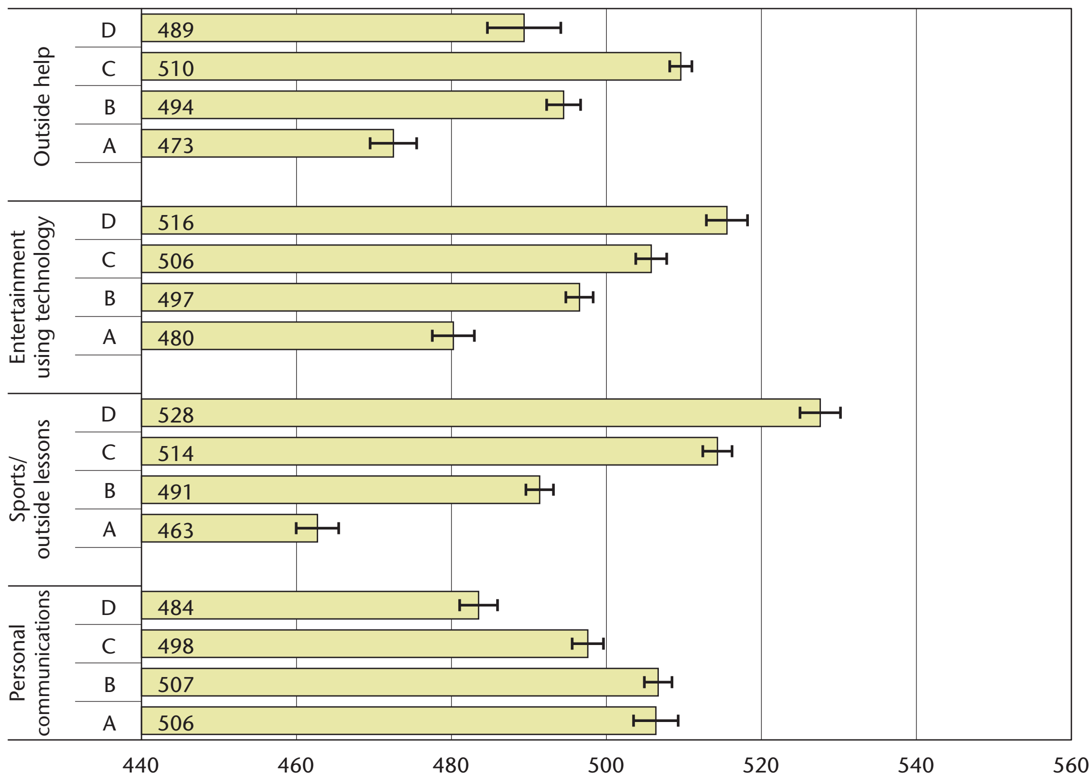


Chart 5-9 gives the mean mathematics scores for the four levels on these four factors. The pattern for entertainment and sports/lessons is clear and linear. The more time students spend on these activities, the lower their mathematics scores. A non-linear pattern emerges for seeking outside help. Students with the lowest level (category D) on this factor have lower scores than those who seek at least some help (those in category C). Beyond this, the pattern is one of decreased scores as the amount of help sought increases. The pattern for personal communications is one of increased scores as this activity increases up to the highest two levels, but remains about the same for these two levels.

Chart 5-9 **Mean mathematics scores for out-of-school activities**



Early mathematics learning

An initial set of seven yes/no items on how students first learned mathematics resolved into two factors, labelled "informal" and "formal" based on the pattern shown in Table 5-3.

A second set of 15 items asked about specific mathematics learning activities engaged in by the student before starting school. These items were on a three-point frequency scale (rarely or never, sometimes, often). This scale gave three factors as shown in Table 5-4. The first is named informal, as it corresponds to some of the items on the previous set. The second is a distinct drill and practice factor, corresponding approximately to the second factor above. The third encompasses items that are less formal than those already called informal, being associated mainly with play.

Table 5-3 **Early mathematics learning methods questionnaire items and factors**

Factors	Items
Informal	• I was taught to count by saying the numbers • I was taught to add by counting • I used materials such as blocks and tiles • I used a lot of diagrams and pictures
Formal	• I was taught to solve word problems • I filled out worksheets • I was taught to memorize the multiplication tables

Table 5-4 **Early mathematics learning activities questionnaire items and factors**

Factors	Items
Informal	• Counted objects • Formed patterns (e.g., beads, tiles) • Sorted objects (e.g., toys) • Compared objects (e.g., sizes) • Identified shapes • Recited numbers • Played board games (e.g., Snakes and Ladders) • Used materials such as blocks and tiles (e.g., Lego)
Drill and practice	• Did problems in a workbook • Was drilled on mathematics facts
Play	• Watched television programs about numbers (e.g., *Sesame Street*) • Played computer mathematics games • Sang number songs • Played board games (e.g., Snakes and Ladders) • Used materials such as blocks and tiles (e.g., Lego) • Played mathematics games (e.g., dominos, cards, dice) • Read books about numbers, shapes, or other mathematical ideas

Charts 5-10 and 5-11 give the mean factor scores associated with the informal and formal learning methods of early mathematics learning factors by population. In general, differences between populations are relatively small. For informal learning methods, none of the populations are above the Canada mean, though several are below. For formal learning methods, only Nova Scotia French is above, while several others are below. Saskatchewan French and Alberta French, along with Newfoundland and Labrador and Nova Scotia English, are below the Canada mean on both factors.

Charts 5-12 to 5-14 show the mean factor scores for the three early mathematics learning activities. For informal activities, Yukon stands out as having a lower mean than any other population except New Brunswick English. All others are close to the Canada mean. For drill and practice, Alberta French, British Columbia French, and Ontario French are significantly above the Canadian average. Six populations, from Manitoba English to Prince Edward Island on the chart, are significantly below the Canadian average. For play activities, Ontario English, Newfoundland and Labrador, and Nova Scotia English are above the Canadian average, and New Brunswick French, Yukon, Quebec French, and Saskatchewan French are below.

Chart 5-10 **Mean factor scores for informal early mathematics learning methods by jurisdiction and language**

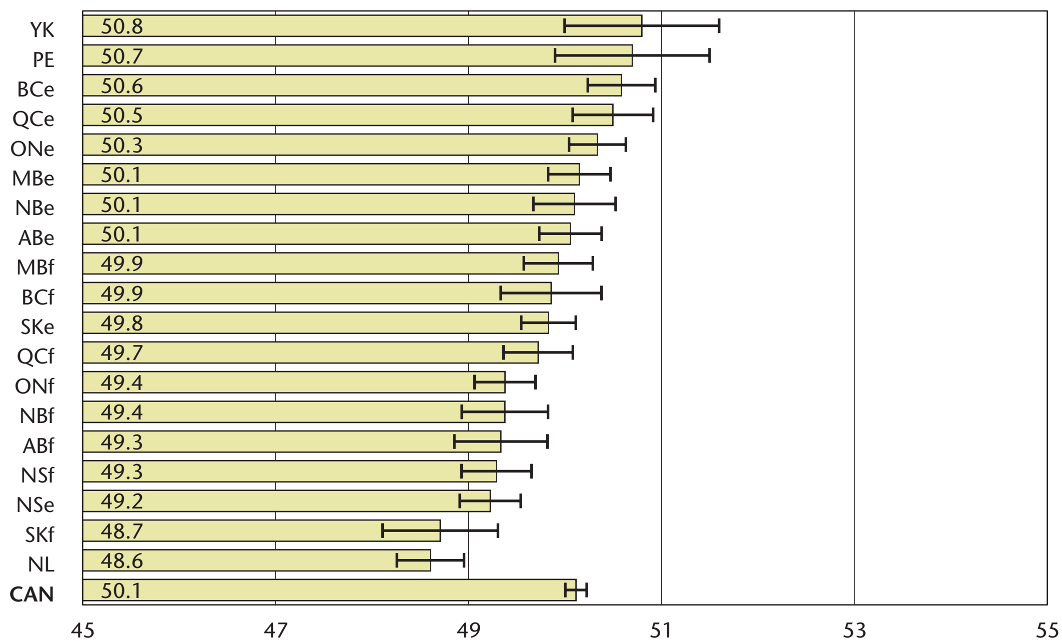


Chart 5-11 **Mean factor scores for formal early mathematics learning methods by jurisdiction and language**

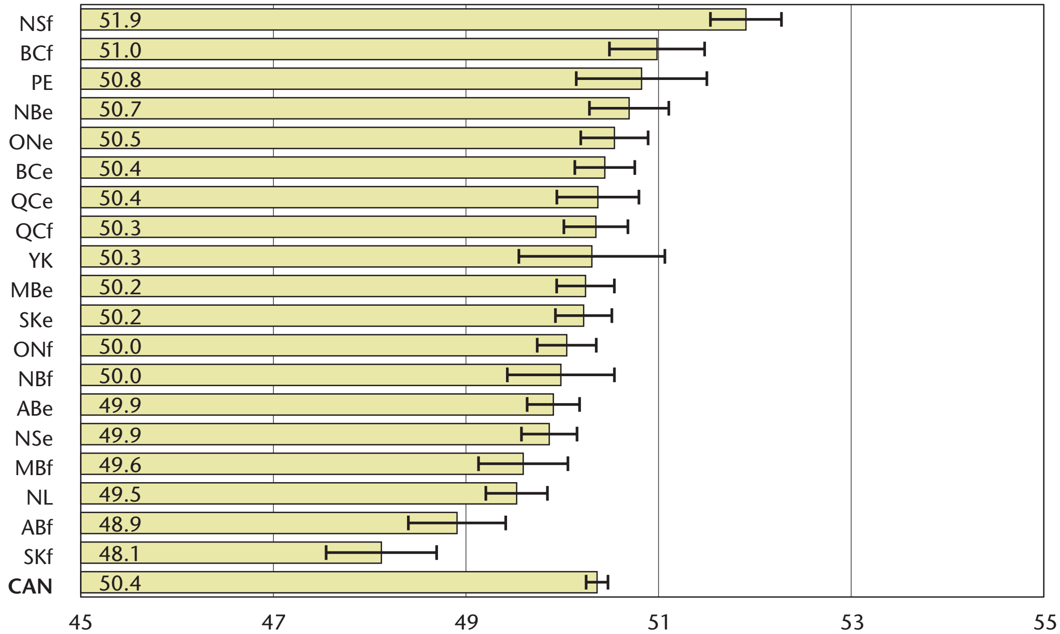


Chart 5-12 **Mean factor scores for informal early mathematics learning activities by jurisdiction and language**

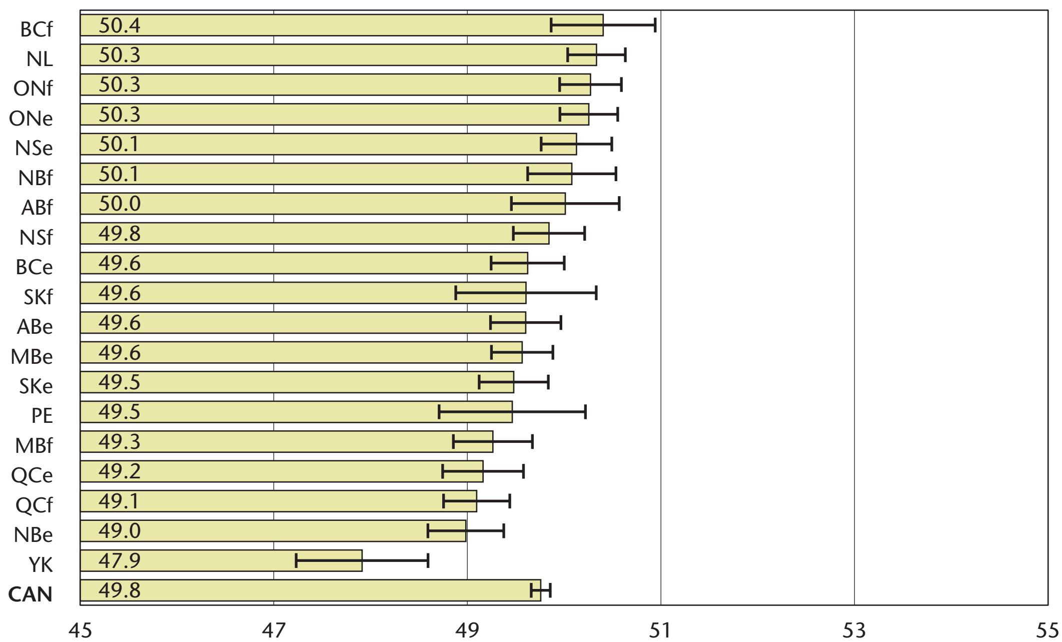


Chart 5-13 **Mean factor scores for drill and practice early mathematics learning activities by jurisdiction and language**

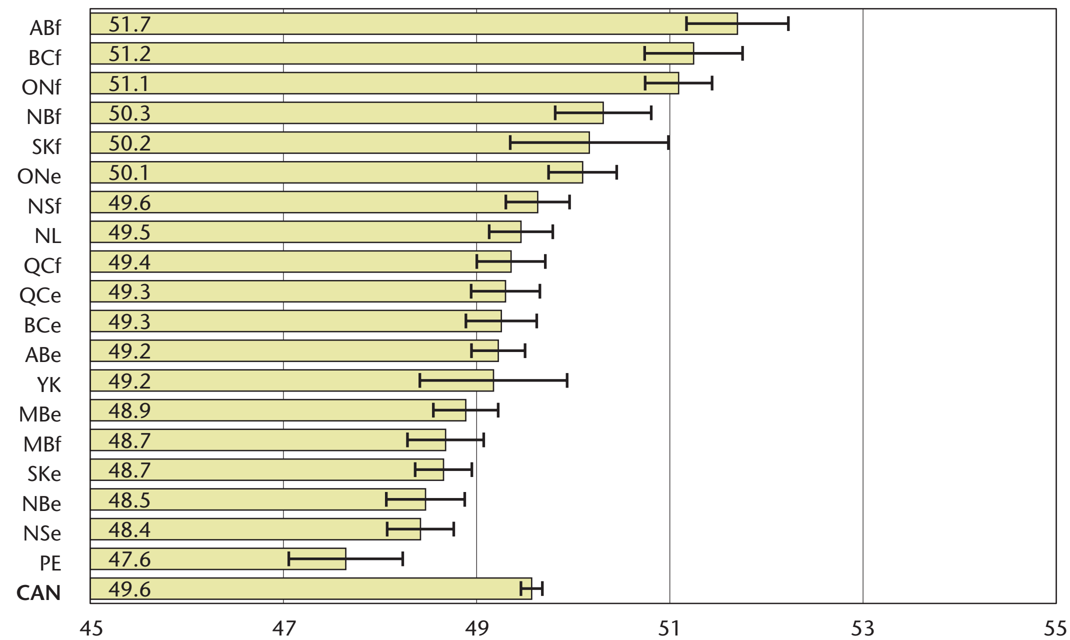
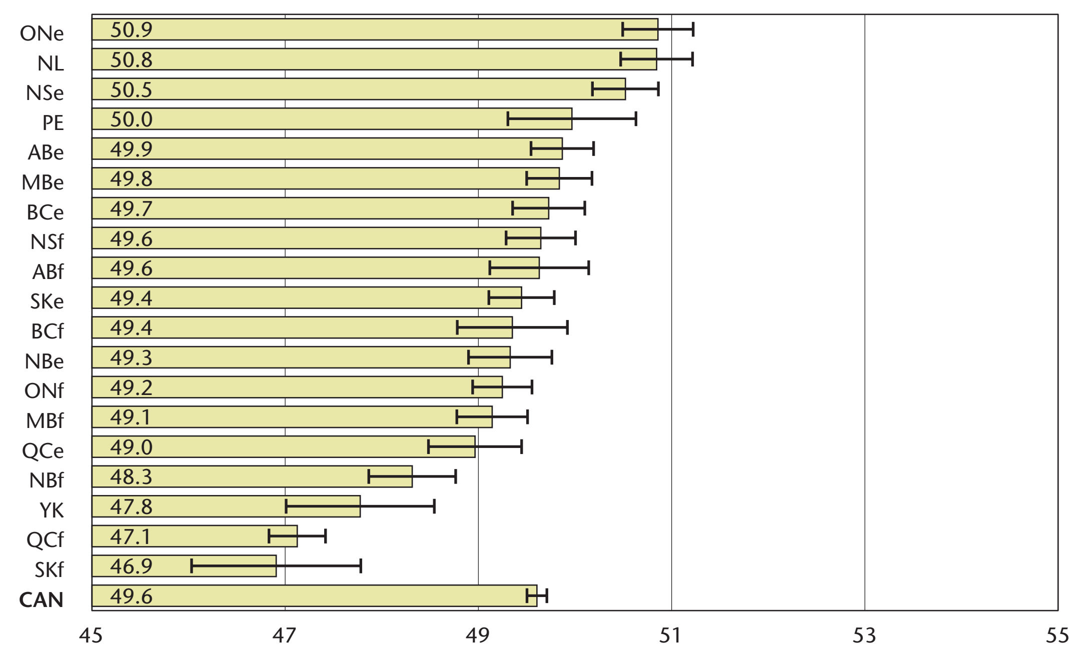


Chart 5-14 **Mean factor scores for play early mathematics learning activities by jurisdiction and language**



Charts 5-15 and 5-16 show the mean mathematics scores for the four levels of each of these factors. Higher levels of informality for both methods and activities are associated with higher mathematics scores. Formal methods show the opposite effect, though the pattern is not linear. Practice also shows a non-linear effect, with those at the two extremes of the scale having higher mathematics scores than those in the two middle categories. Finally, play shows a pattern similar to informal activities, with higher scores for higher levels on the play scale. Although separate in the factor analysis, the questionnaire items indicate that play is more closely aligned to informal than to formal activities.

CHART 5-15 **Mean mathematics scores by early mathematics learning methods**

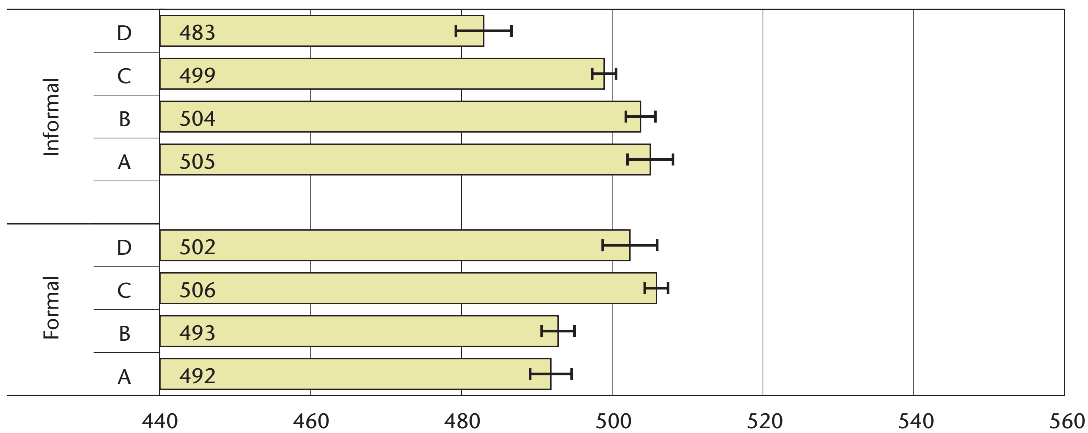
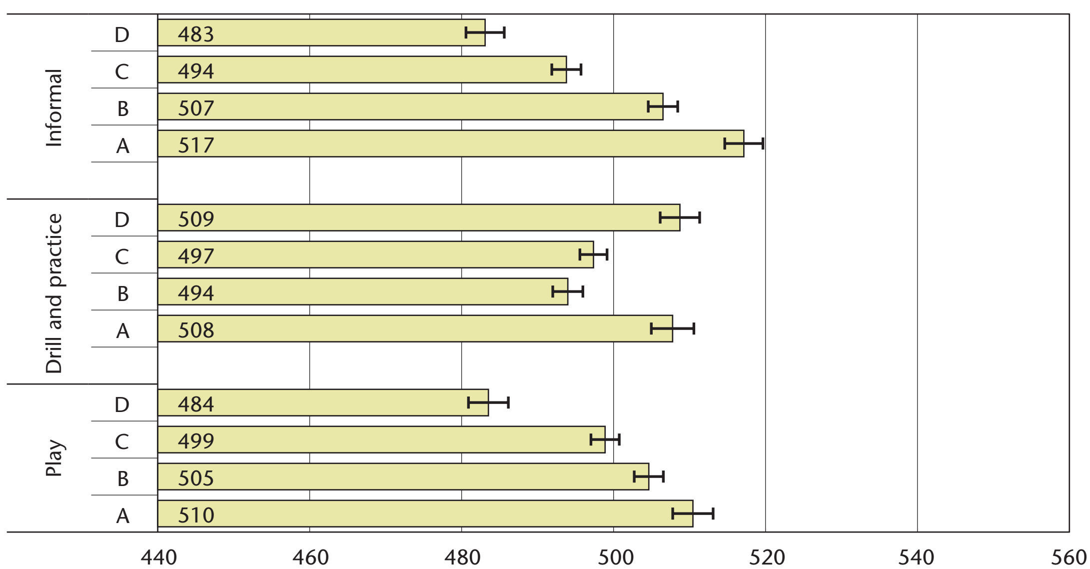


CHART 5-16 **Mean mathematics scores by early mathematics learning activities**



Mathematics learning strategies

A 16-item frequency scale was used to determine the strategies students use in mathematics learning (Table 5-5). A four-factor solution emerged, with clear distinctions on multiple items. The first factor is labelled "graphics/pictorial" because the highest loadings had to do with using pictures, graphs, or diagrams. The second factor is more difficult to label but is called "learning techniques" because the actions are more concrete than those for the third factor, labelled "strategic approach." The second factor includes a couple of items associated with "persistence," a factor which has emerged for other scales. However, the persistence items do not have the highest loadings, so a more generic label — "learning techniques" — was chosen. The factor labelled "strategic approach" includes items often taught as generic ways of solving new mathematics problems. Finally, a factor related to seeking external sources of help also emerged. This is labelled "Internet/tutor" after the two items loading on this factor. This factor appears to resemble that labelled "outside help" in Table 5-2. However, these two factors are not highly correlated.

Table 5-5 **Mathematics learning strategies questionnaire items and factors**

Factors	Items
Graphics/pictorial	• Create a diagram or picture • Draw a table, chart, or graph • Underline key words • Model with concrete materials • Look for examples in textbook or notes
Learning techniques	• Use a calculator • Ask for help • Keep trying • Re-read the problem • Find a quiet place to work
Strategic approach	• Work backwards • Look for patterns • Use friendly numbers • Guess and check
Internet/Tutor	• Use the Internet • Work with a mathematics tutor

Charts 5-17 to 5-20 show the mean factor scores on these factors by population. On the graphics/pictorial factor, Newfoundland and Labrador, Quebec English, and Ontario English stand out as having means both above the Canada mean and higher than most other populations.

Several populations have means below the Canada mean. For the learning techniques factor, Quebec English and Quebec French are higher than the Canada mean and higher than all other populations, with most others being below the Canada mean.

The strategic approach factor shows Ontario French, British Columbia French, and Ontario English above the Canada average and above most other populations, while New Brunswick English and Saskatchewan English are below the Canada average and below most others.

Finally, on the Internet/tutor factor, Newfoundland and Labrador, Quebec English, British Columbia English, and New Brunswick French are above the Canadian average. Four of the francophone populations, along with Saskatchewan English and Manitoba English, form a cluster at the low end of the distribution, different from all others except Yukon.

Chart 5-17 **Mean factor scores for the mathematics learning strategies factor "graphics/ pictorial" by jurisdiction and language**

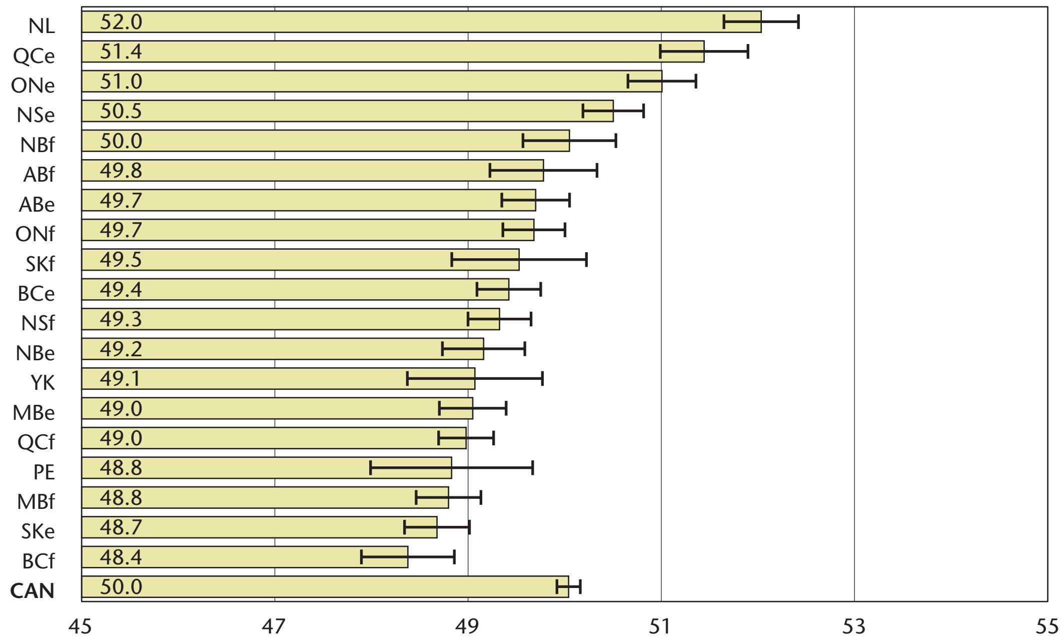


Chart 5-18 **Mean factor scores for the mathematics learning strategies factor "learning techniques" by jurisdiction and language**

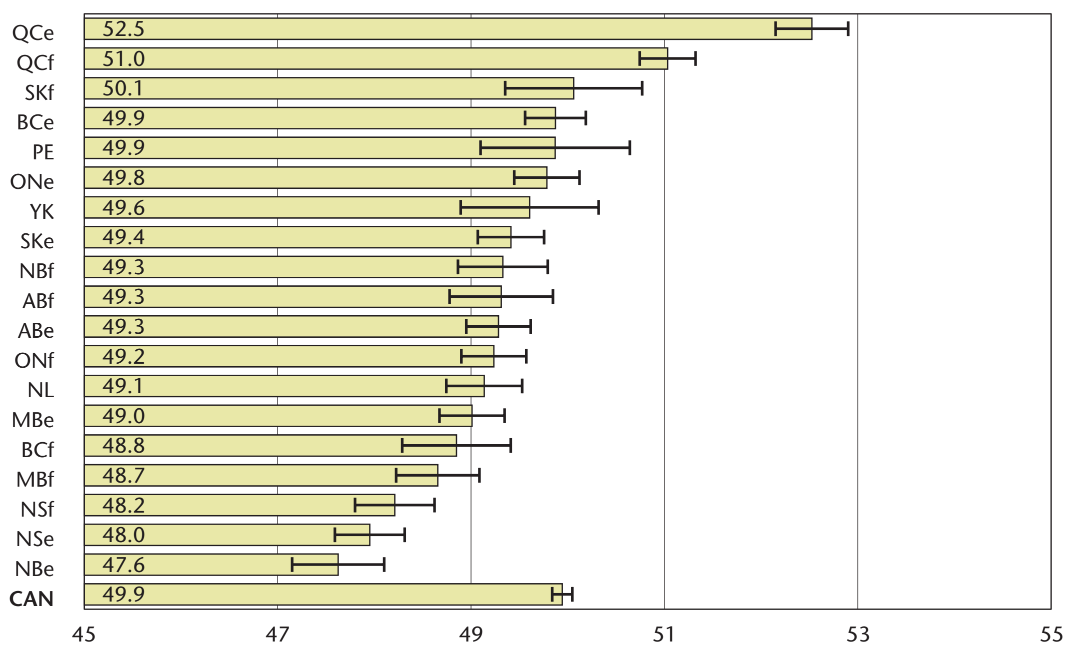


Chart 5-19 **Mean factor scores for the mathematics learning strategies factor "strategic approach" by jurisdiction and language**

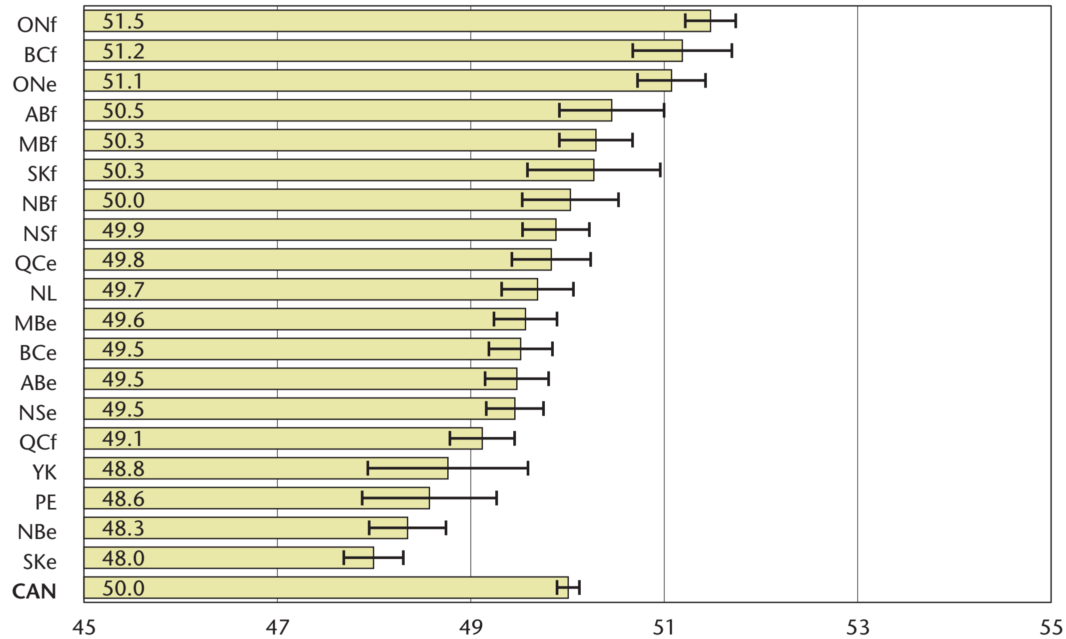


Chart 5-20 **Mean factor scores for the mathematics learning strategies factor "Internet/ tutor" by jurisdiction and language**

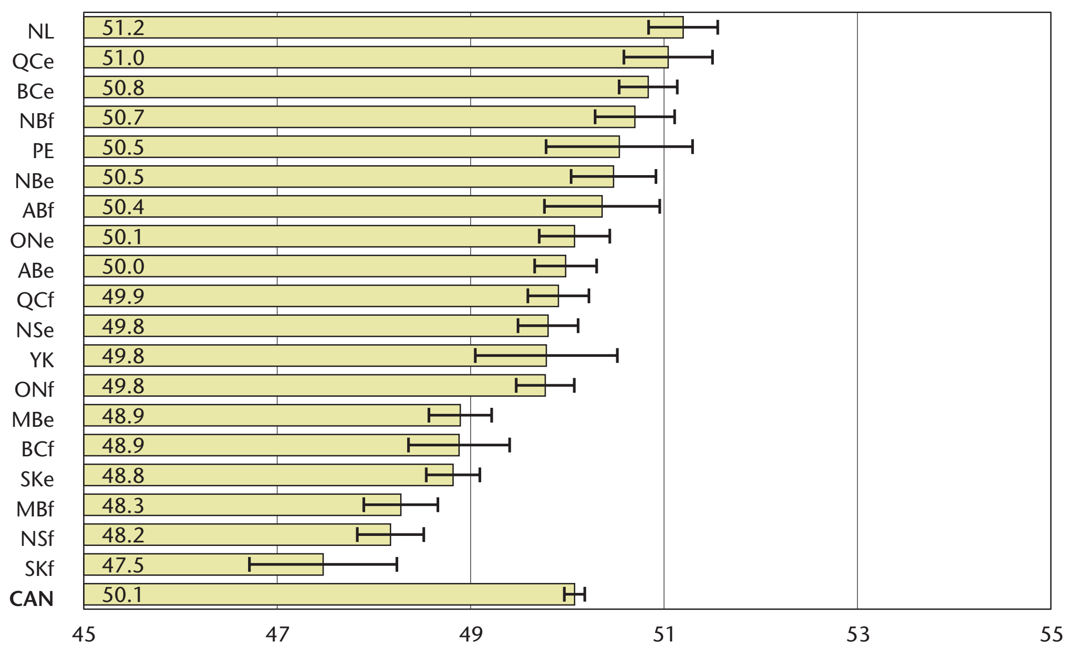
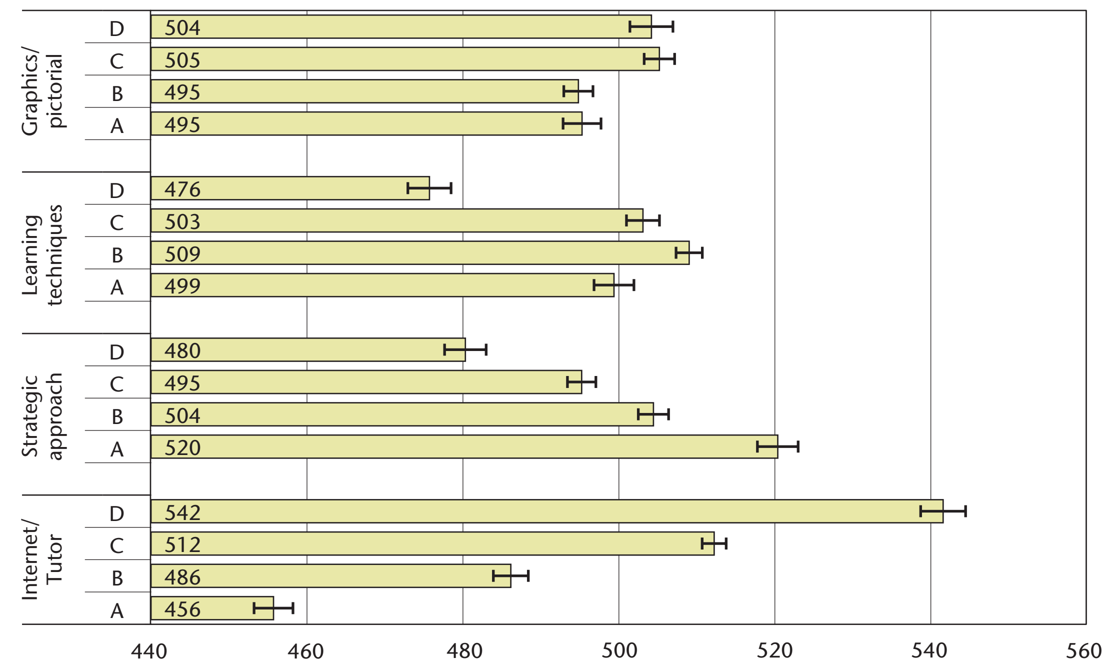


Chart 5-21 shows the distributions of mean mathematics scores across the four levels of these learning strategy factors. The graphics/pictorial factor shows a pattern of higher mean scores for the lowest two levels and lower mean scores for the highest two levels. The learning techniques factor shows a non-linear pattern. Those at the lowest level of this factor have a significantly lower mean score than those at the other three levels, while those at the top (level A) are significantly lower than those at the next level (level B). The strategic approach factor shows a more distinct linear pattern of higher mean scores for higher levels of this factor. The opposite is true, and even more pronounced, for the Internet/tutor factor, with lower mean scores associated with higher levels on the factor.

Chart 5-21 **Mean mathematics scores by mathematics learning strategies**



Multiple regression effects

The effects of mathematics behaviours and strategies were modelled using two-level simple and multiple regression models as before. Again, the simple regression coefficients represent absolute (uncontrolled) effects, and the multiple regression coefficients represent the relative effect of each variable controlling for all others in this set of variables. Since all of the variables in this group are based on factor scores, the coefficients are comparable across variables as well as across the simple and multiple regression effects. The coefficient in each case is interpreted as the change in mathematics score associated with one standard deviation change in the factor score.

Chart 5-22 gives the simple and multiple coefficients for student approaches to difficult mathematics problems. All of these are statistically significant in both models. There is no significant difference between the two sets of coefficients, indicating that these three factors act independently of each other in influencing achievement. Persistence has a strongly positive influence on achievement, while the use of on-line help and seeking help from others both show negative but somewhat weaker effects.

CHART 5-22 **Regression coefficients for encountering difficult mathematics problems**

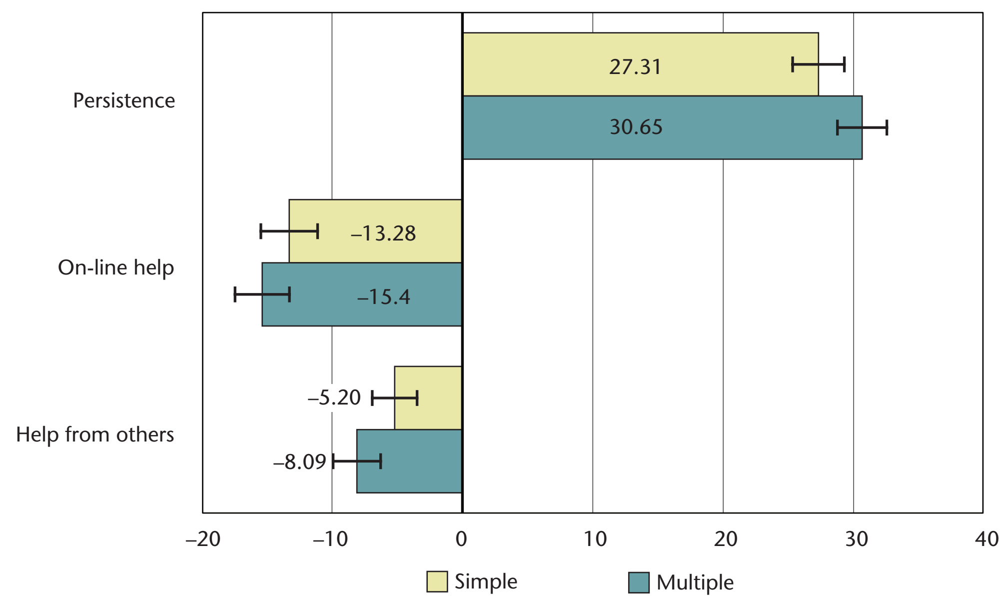


Chart 5-23 presents the results for time on out-of-class activities. Again, all of these effects are statistically significant in both the simple and multiple regression models but with no significant difference between the two models. Spending time seeking outside help with school work, on entertainment using technology, and on sports or other lessons outside of school has a negative effect on achievement. Spending time on personal communications (e.g., using the telephone, texting, Internet, e-mail, or playing mathematics-related computer games) has a positive effect.

Chart 5-23 **Regression coefficients for out-of-class activities**

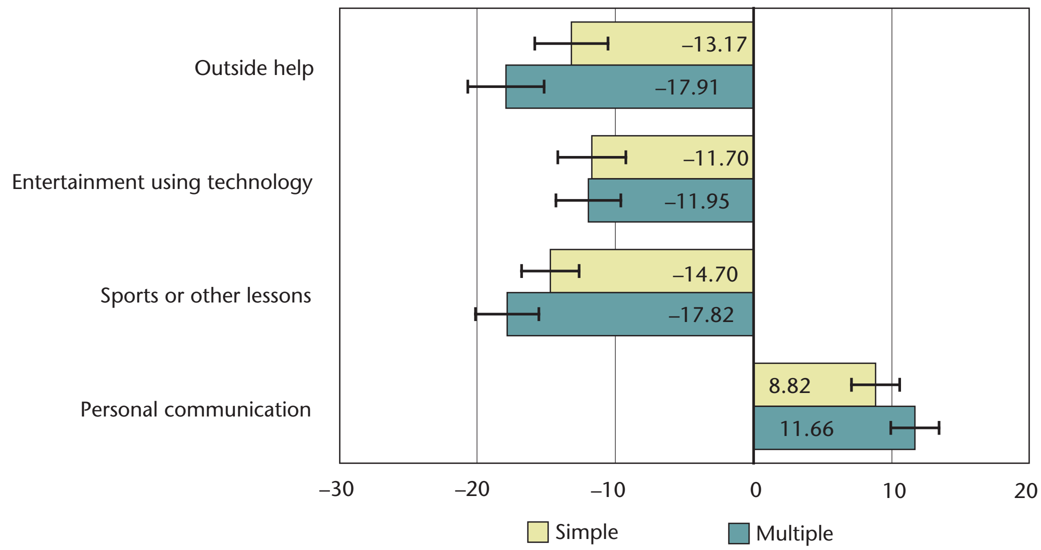
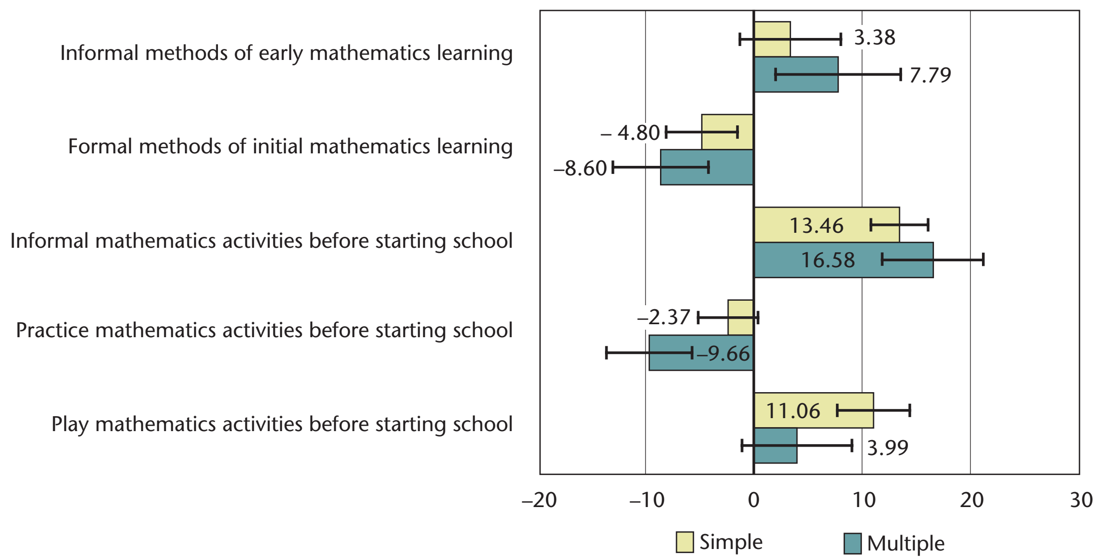


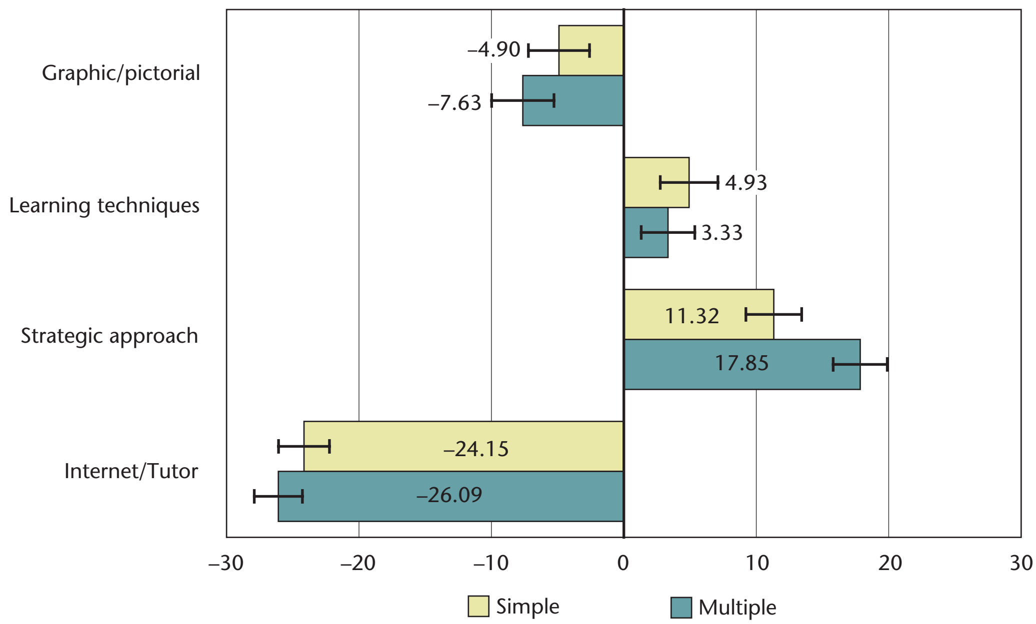
Chart 5-24 gives the effects of early mathematics learning methods and activities. The effect of using informal methods (e.g., counting, saying numbers, using blocks and tiles) is not significant in the simple regression model but becomes significant when the other factors in this set are controlled. Using more formal methods (e.g., filling out worksheets, memorizing multiplication tables) has significant negative effects in both models. Informal mathematics activities (e.g., counting, sorting, comparing, patterning) shows a significant positive effect for both models. The effect of drill and practice activities is non-significant in the simple regression model but becomes significantly negative in the multiple regression model. The opposite is true for play activities, for which the significantly positive effect in the simple regression model becomes non-significant when other variables are controlled.

Chart 5-24 **Regression coefficients for early mathematics learning methods and activities**



Finally, the coefficients for mathematics learning strategies given in Chart 5-25 are consistently significant in both models. Use of graphical/pictorial strategies has a small but significant negative effect, whereas use of learning "techniques" (e.g., calculator, asking for help, rereading the problem) has a small but significant positive effect. Taking a strategic approach to mathematics learning (e.g., working backwards, looking for patterns, using friendly numbers) has a positive effect, which increases significantly when other variables in this set are controlled. Overall, the strongest, negative, effect is for seeking help, a result consistent with that found for earlier similar factors. It might be expected that weaker students would be the ones seeking help, and it might be the case that such help does improve the performance of such students. However, these results show clearly that seeking help is not a transformative strategy in changing generally low into generally high performance.

CHART 5-25 **Regression coefficients for mathematics learning strategies**



6 INSTRUCTIONAL CLIMATE

Instructional climate refers to features of the school and classroom climate that can be expected to have a bearing on mathematics achievement. Relevant aspects include the school's overall philosophy and areas of emphasis in mathematics, class size, influences on decision making, and the presence of special-needs students. Most of the data in this area came from the teacher and school questionnaires.

Areas of emphasis in mathematics

Principals were asked to rate the degree of emphasis on eight aspects of mathematics teaching and learning in their schools, on a three-point scale from "little or no emphasis" to "a lot of emphasis." This scale yielded three factors, as shown in Table 6-1. The first two factors distinguish between an emphasis on "generic skills" and "basic skills." The third factor encompasses emphasis on performance on external assessments.

TABLE 6-1 **Areas of emphasis in mathematics items and factors**

Factors	How much do you emphasize the following in teaching mathematics in your school?
Generic skills	• Having students perform to the best of their abilities • Using a variety of strategies to challenge each student • Developing the well-rounded individual • Understanding concepts and big ideas
Basic skills	• Basic mathematical skills • Computational skills
Performance	• The knowledge and understanding needed for students to do well on pan-Canadian or international assessments • The knowledge and understanding needed for students to do well on provincial/territorial assessments

Charts 6-1 to 6-3 give the mean responses across populations on these three factors. For generic skills, none of the populations are above the Canadian average. Francophone populations tend to be at the lower end of the distribution of this variable, though not all are significantly below the Canadian average.

Effects tend to be in the opposite direction, and are more pronounced, for basic skills where most of the populations, including all of the francophone populations, are above the Canadian average. No population is below the Canadian average on this variable.

The results for emphasis on performance on external assessments show still greater differentiation across populations. Three English populations — Manitoba, British Columbia, and Saskatchewan — are above the Canadian average, while five populations, from Newfoundland and Labrador to New Brunswick French on the chart, are below.

CHART 6-1 **Mean factor scores for emphasis on generic skills by jurisdiction and language**

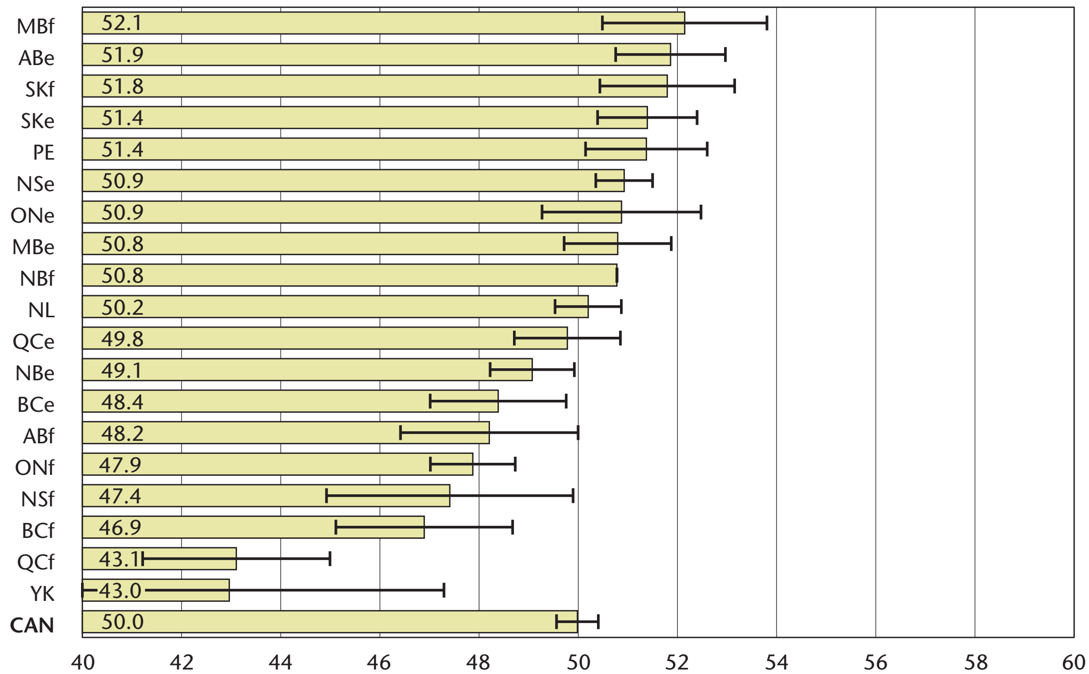


CHART 6-2 **Mean factor scores for emphasis on basic skills by jurisdiction and language**

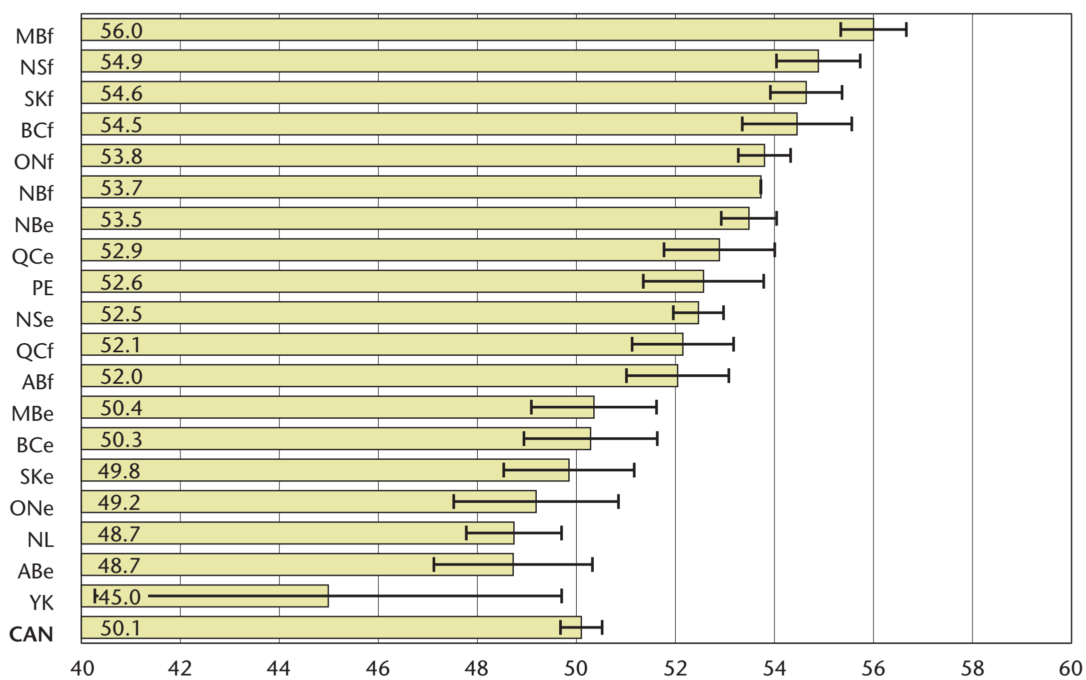


Chart 6-3 **Mean factor scores for emphasis on performance on external examinations by jurisdiction and language**

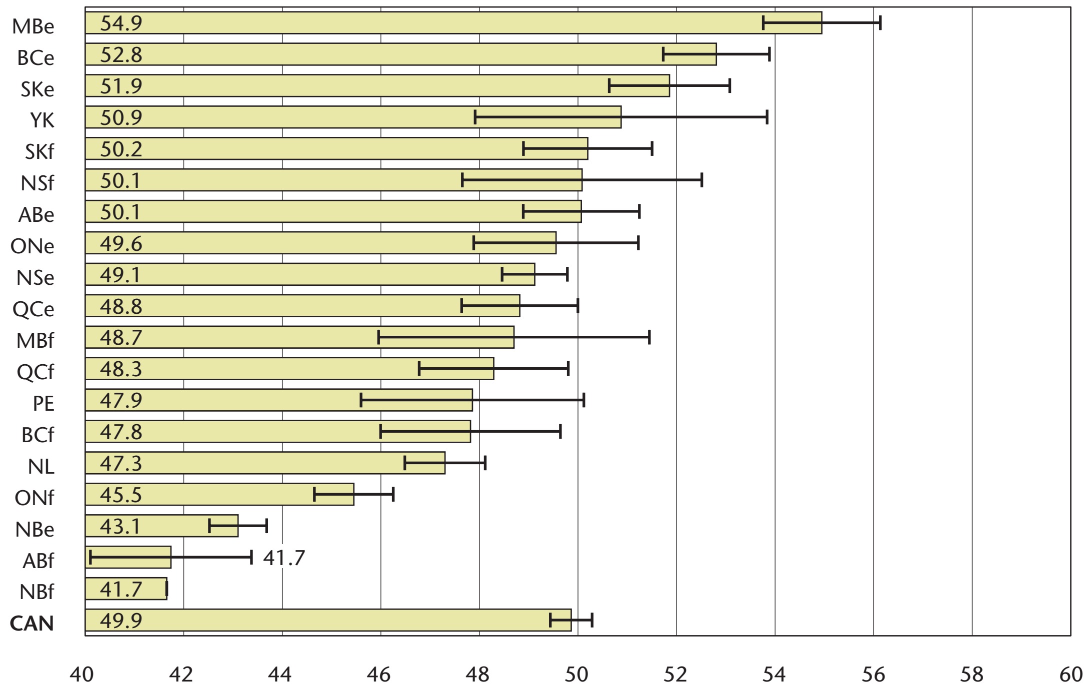
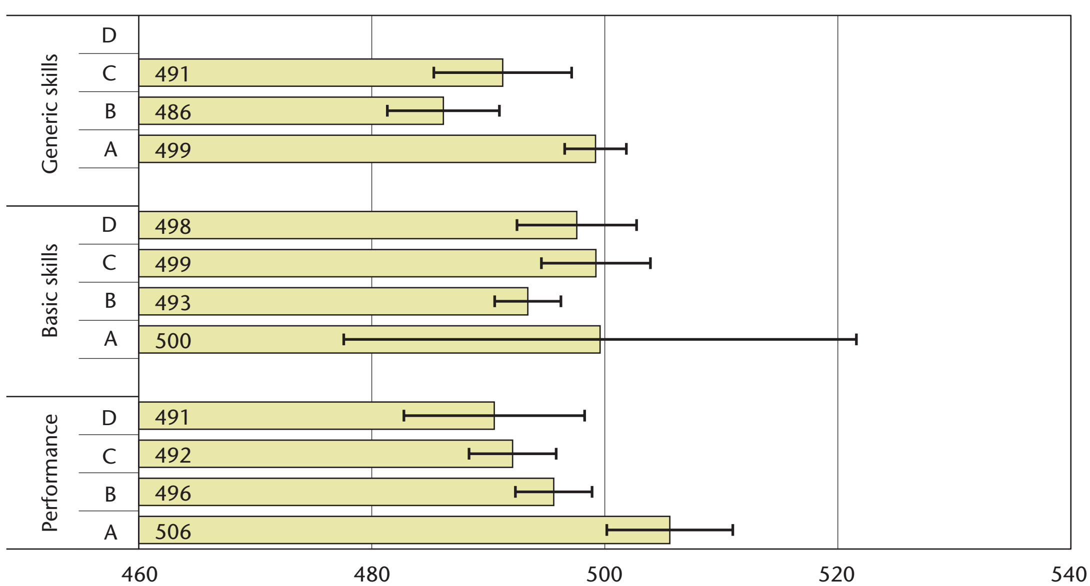


Chart 6-4 gives the change in mathematics mean score for each standard deviation unit on each of these three factors.[16, 17] For emphasis on generic and basic skills, there is no clear trend. However, for emphasis on performance on external assessments, schools that are above one standard deviation unit (category A) have significantly higher mean scores than those at other levels on this factor.

Chart 6-4 **Mean mathematics score by areas of emphasis in mathematics**



[16] Readers are reminded that the four categories in the charts represent units of one standard deviation above or below the mean on the factor: D = below –1 SD, C = –1 to 0 SD, B = 0 to +1 SD, and A = above +1 SD.

[17] There are no schools in Category D for generic skills.

Class size

Class size information was obtained by asking teachers to report the average number of students in their mathematics classes. Chart 6-5 shows class size ranges by jurisdiction and language. The most striking thing about these distributions is the extent of variation both within and between populations. Quebec French has the largest percentage of classes with 30 or more students. Several populations have 10 per cent or fewer of their classes in the highest range. With the exception of Quebec French, most of the populations with the smallest class sizes are francophone.

Chart 6-6 gives mean mathematics scores for teachers with class sizes in the various ranges. These results show that performance is related to class size in the opposite direction from what is generally expected. In terms of statistical significance, three groups can be seen from the chart. The lowest performance is found in the smallest classes. The three largest class sizes — from “20 to 24” to “30 or more” — show significantly increasing achievement with increased class size.

Although these results are counterintuitive, and inconsistent with some experimental studies of class size, they are consistent with what has been previously found in SAIP and PISA studies, and for reading in PCAP 2007. Class size, like many other variables in this analysis, may be confounded with many other factors, particularly school size and location. It is important to examine the class size effect with such variables controlled. This is done as part of the multiple regression analysis to be presented at the end of this chapter.

Chart 6-5 **Teacher reported mathematics class size ranges by jurisdiction and language**

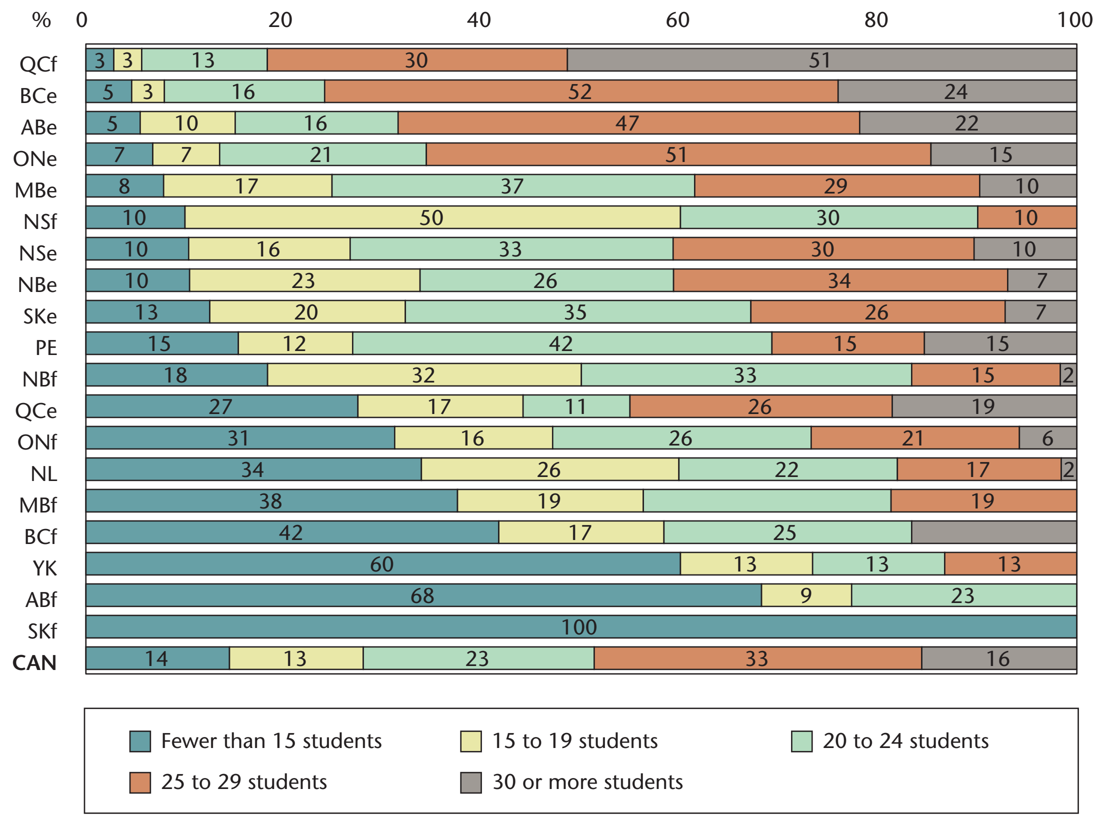
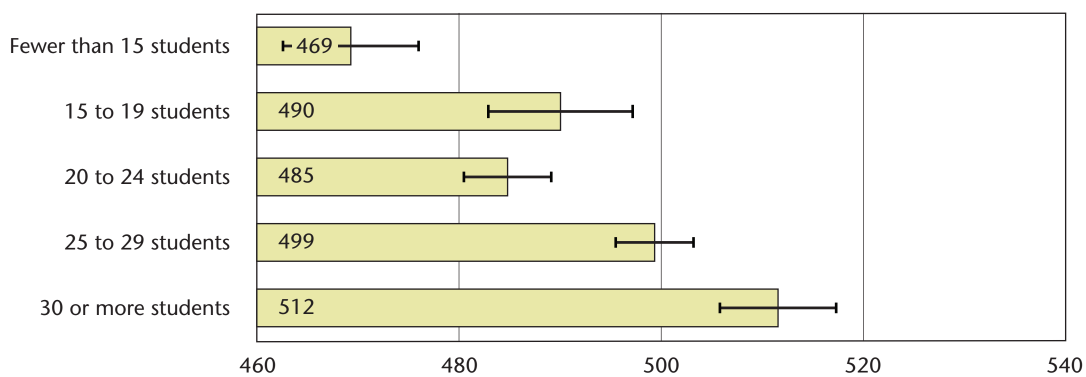


Chart 6-6 **Teacher mean mathematics scores by class size ranges**



Sources of influence on school programs

Principals were given a series of 15 questions about the extent to which various people or agencies influence decisions about school programs and activities. The scale was a four point one from “not at all” to “a lot.” Factor analysis of this question set yielded four factors, as identified in Table 6-2. Two of these factors, labelled “external” and “internal,” may reflect a more general underlying trait called “school autonomy” that has frequently been referenced in the literature on school improvement. The remaining two factors may be interpreted in the same way, with assessment being a source of external influence and students/parents being a source of internal influence.

Table 6-2 **Sources of influence on school programs items and factors**

Factors	Items
External	• Church or religious groups • Textbooks and textbook publishers • Access to resources • Teacher groups external to the school • External agencies (e.g., business community)
Internal (–)	• Provincial/territorial curriculum • Individual teachers • Results from classroom assessments • Teachers within departments or subject groups
External assessment	• Assessment results from PISA and PCAP • Provincial/territorial assessment results that do not count toward students’ final marks • Provincial/territorial assessment results that count toward students’ final marks
Students/parents (–)	• Students’ voice or representation • Parent/guardian advisory committees or school councils • Characteristics of the student body
Note: (–) indicates factors that showed negative loadings of their items during the factor analysis.	

Charts 6-7 to 6-10 show mean factor scores on these four factors by jurisdiction and language. Five populations — Saskatchewan and Nova Scotia French, Alberta French and English, and Newfoundland and Labrador — are above the Canadian average on the external influence factor. Five are below the Canadian average, including New Brunswick English and French, Quebec English and French, and British Columbia English. This suggests that external influence may be linked to characteristics of particular jurisdictions more than to language groups.

Chart 6-7 **Mean factor scores for external influence by jurisdiction and language**

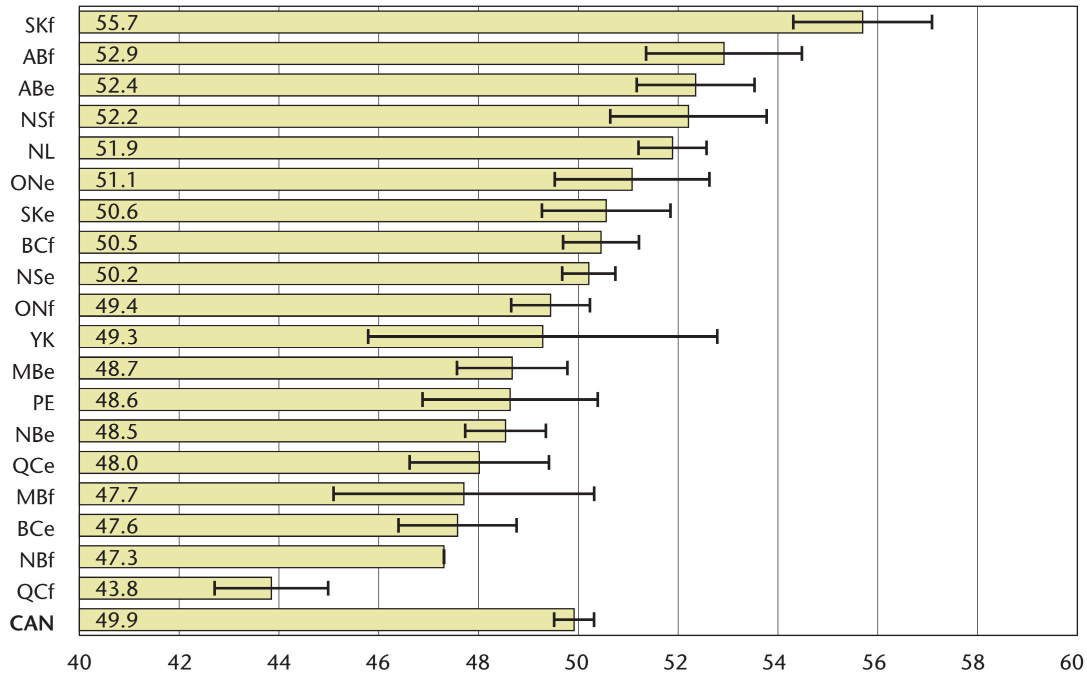


Chart 6-8 **Mean factor scores for internal influence by jurisdiction and language**

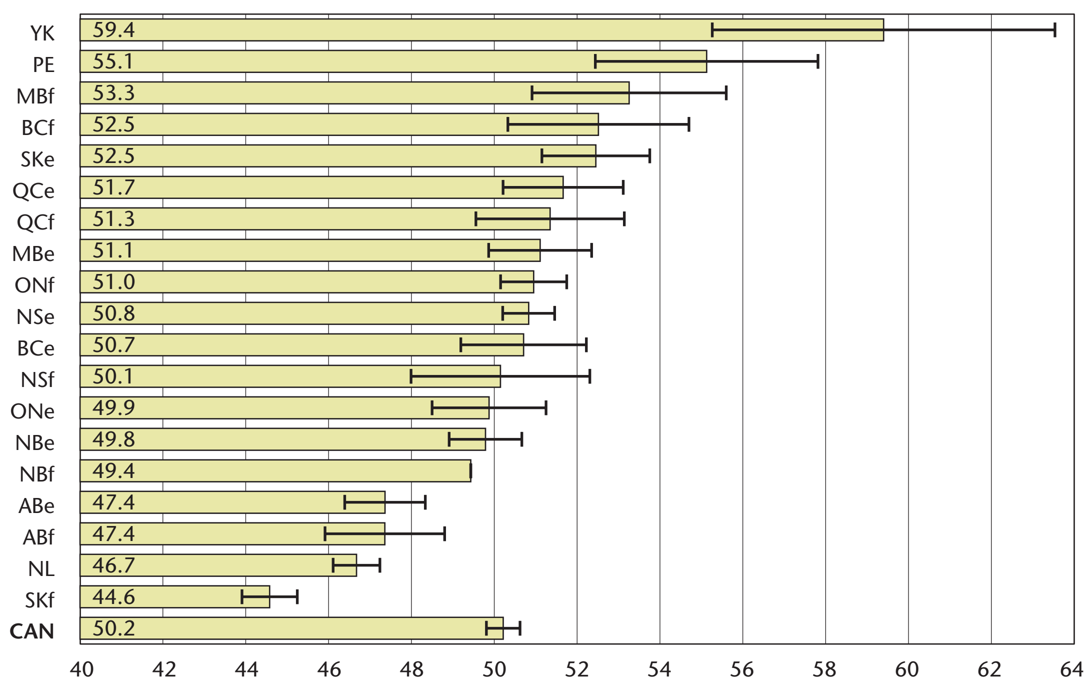


CHART 6-9 **Mean factor scores for external assessment influence by jurisdiction and language**

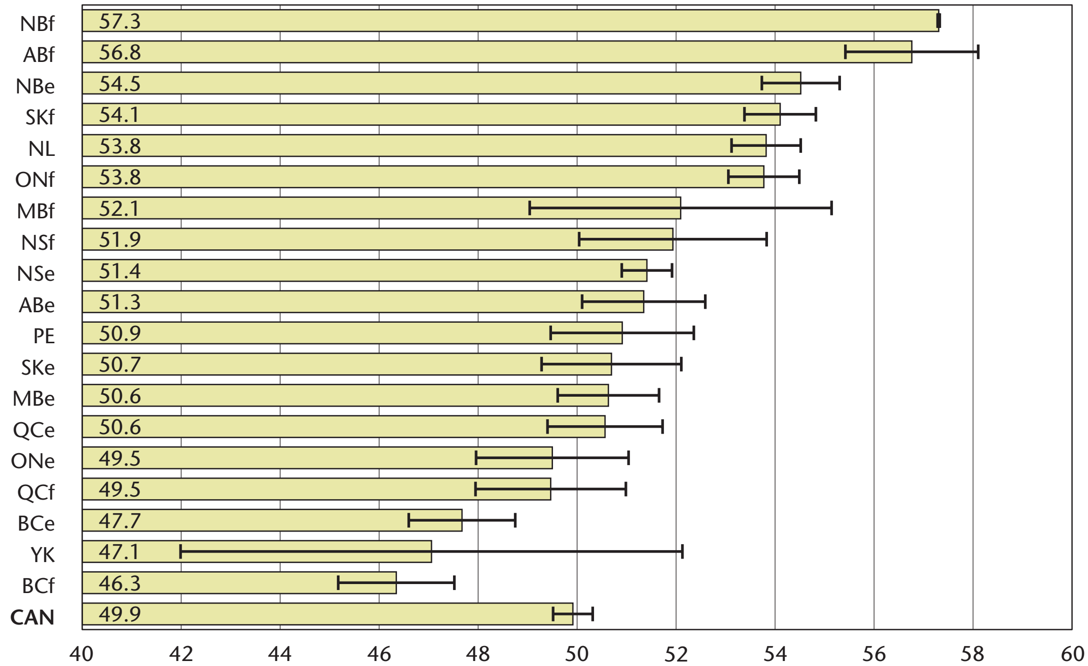
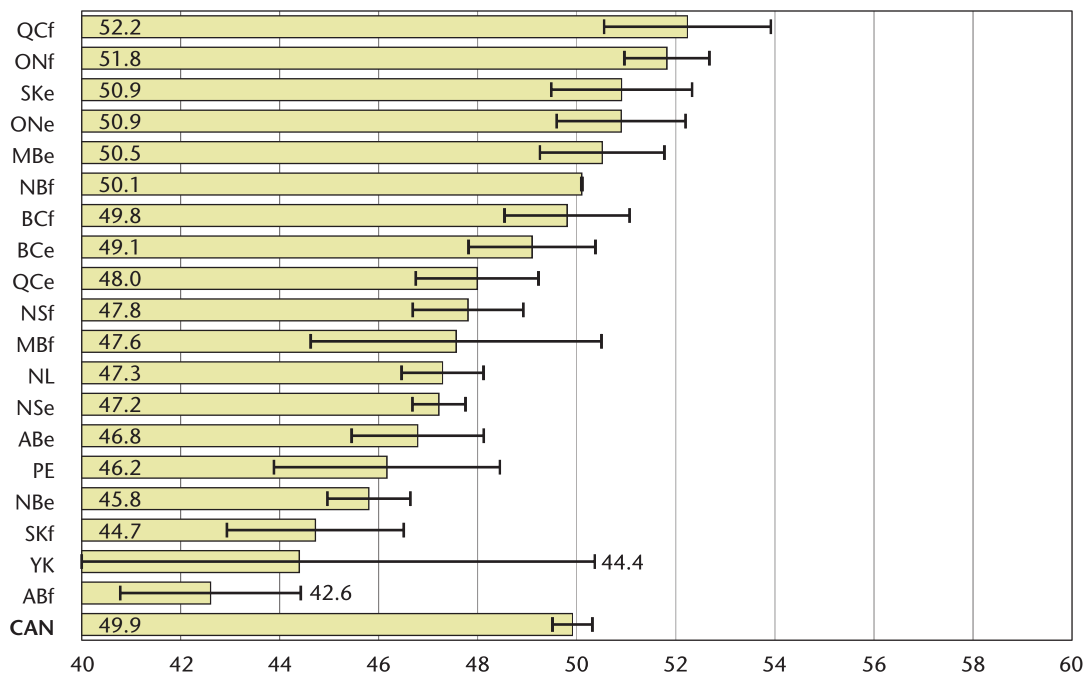


CHART 6-10 **Mean factor scores for student/parent influence by jurisdiction and language**



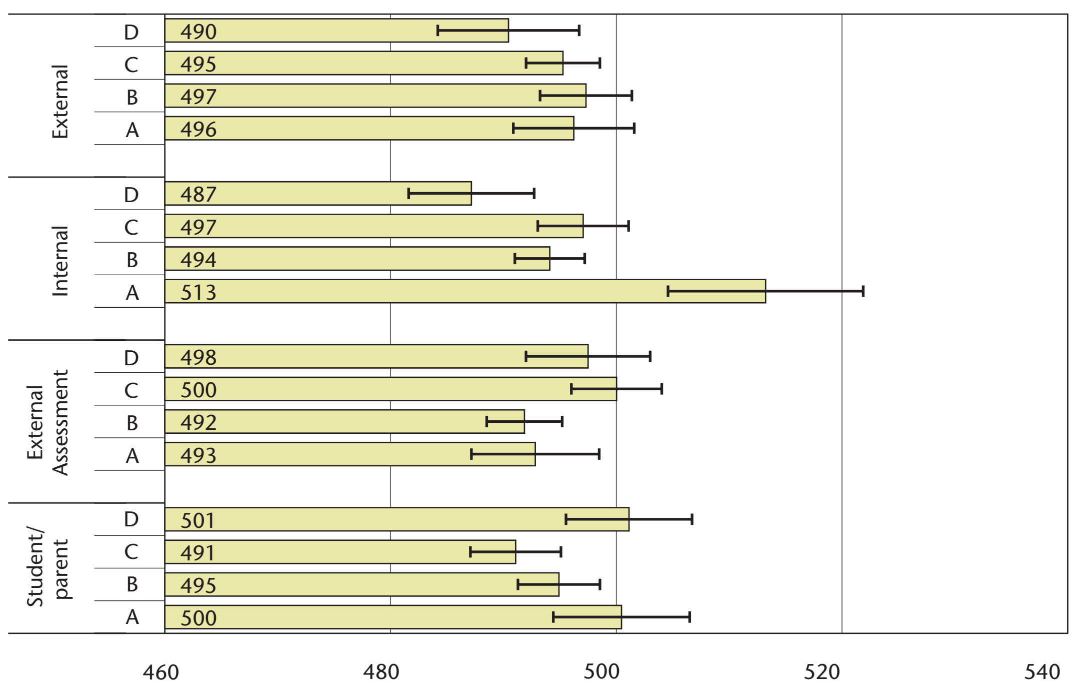
For the internal influence factor, the spread is also quite wide. Yukon, Prince Edward Island, Manitoba French, and Saskatchewan English are above the Canadian average, even with their wide confidence intervals. At the low end are Alberta English and French, Newfoundland and Labrador, and Saskatchewan French. The latter populations were at the high end of the external influence scale. This reinforces the intuitive concept that external and internal influence may represent opposite ends of a continuum.

External assessment influence (Chart 6-9) shows three population clusters. Five populations, from New Brunswick French to Ontario French, have factor scores significantly above the Canadian mean and above most other populations. Nova Scotia English is also above the Canadian average. British Columbia, both English and French, are at the low end and are also significantly below most other populations.

Finally, for the student/parent influence factor (Chart 6-10), only two populations — Quebec French and Ontario French — are above the Canadian average which indicates a high level of influence of this factor, while most others are below.

Chart 6-11 shows mean mathematics scores for schools at each of the standard deviation ranges on each of these factors. Only the internal factor shows a significant pattern, with higher internal influence being associated with higher scores.

Chart 6-11 **School mean mathematics scores for sources of influence on school programs**



Presence and accommodation of special-needs students

The school questionnaire included two questions on the placement of special-needs students in the school. The first asked how special-needs students are actually placed and should be placed within the school. Chart 6-12 shows the distribution of preferences for the three placement options across jurisdictions and languages.

In this case, more than one response could be given. The chart thus shows the responses to each option shown cumulatively with values to a maximum of 300 for the three options. This gives an indication of the variety of placement strategies used. For example, almost all Nova Scotia French and British Columbia French schools use all three placement methods. In most other populations, placement in regular classes, with additional adults assigned specifically to the special-needs students, is the most common form of placement; placement in regular classes with the classroom teacher alone is also found in many schools. Only Yukon seems to use placement with the classroom teacher and other adults almost exclusively (keeping in mind that the Yukon sample was only 10 schools).

Chart 6-12 **Placement of mathematics special-needs students by jurisdiction and language**

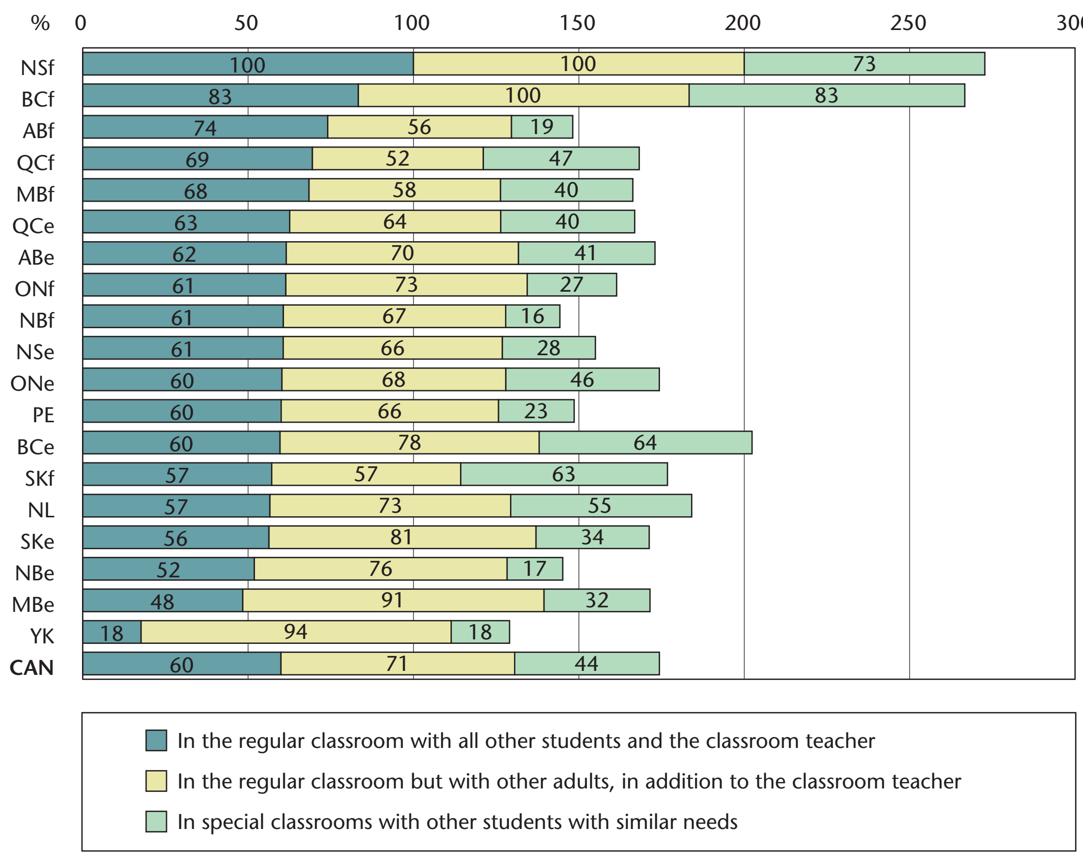


Chart 6-13 shows principals' responses to the question of how special-needs students should be placed for instruction in mathematics. A majority in all populations prefer placement in regular classes with adults other than the regular classroom teacher. The percentage reporting a preference for placement with only the regular teacher varies considerably by population. The percentages favouring placement in special classes is generally lower than for the other two categories. Quebec English and French, along with British Columbia French, show the highest preference for this option.

Finally, principals were asked about the effect on regular mathematics classes of needing to attend to special-needs students. Chart 6-14 gives these results. Again, there is wide variation, especially in the percentages indicating "a lot" of effect. There is a distinct division by language on this question, with more francophone than anglophone principals reporting "a lot" of effect.

Differences in responses to these questions have no significant effect on school mathematics scores, as shown in Chart 6-15.

Chart 6-13 **Principals' perceptions of how special-needs students should be placed for mathematics instruction by jurisdiction and language**

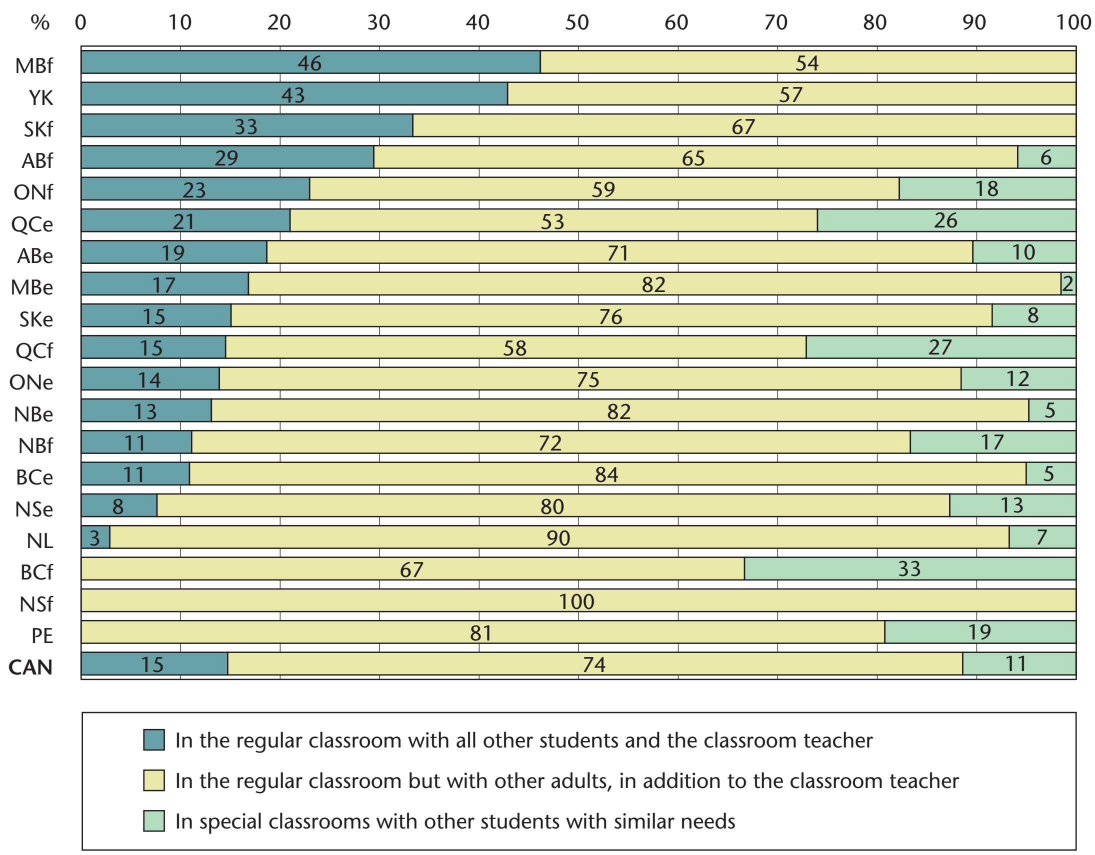


Chart 6-14 **Principals' perceptions of the effect on mathematics classes of having to attend to special-needs students by jurisdiction and language**

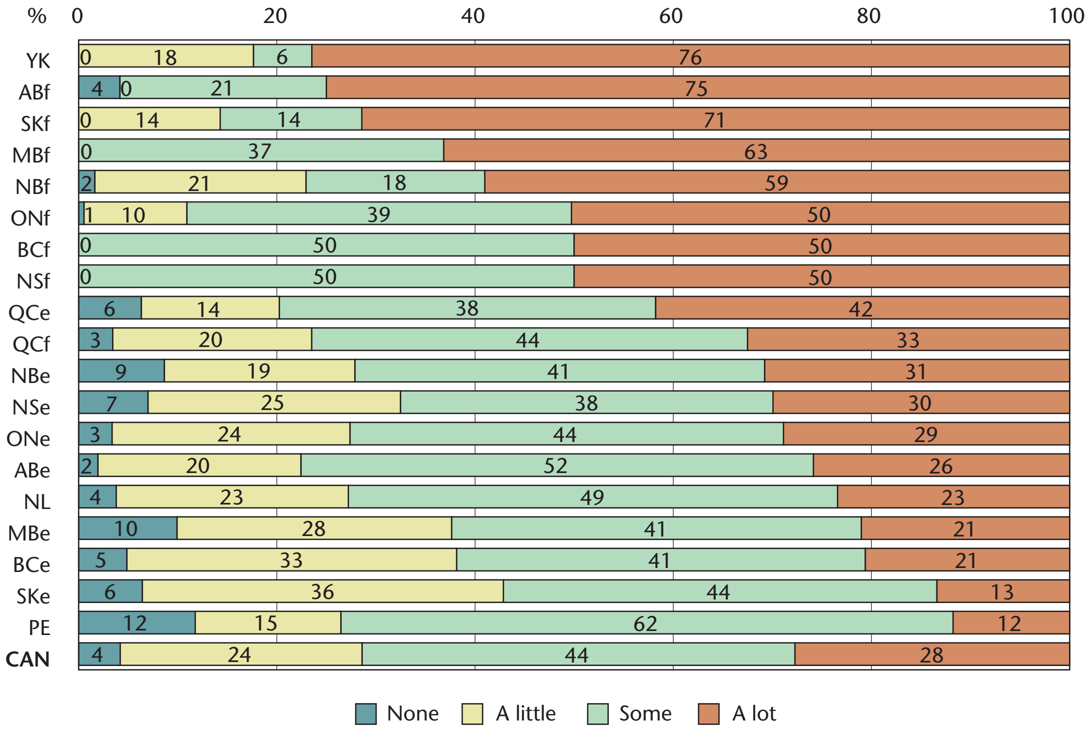
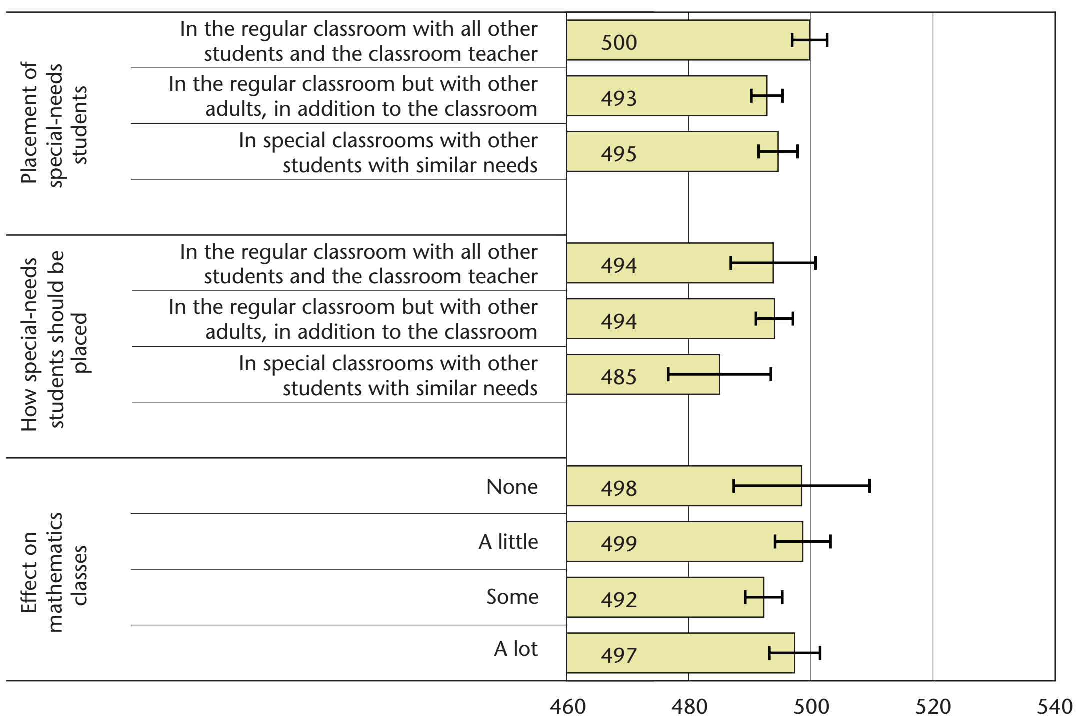


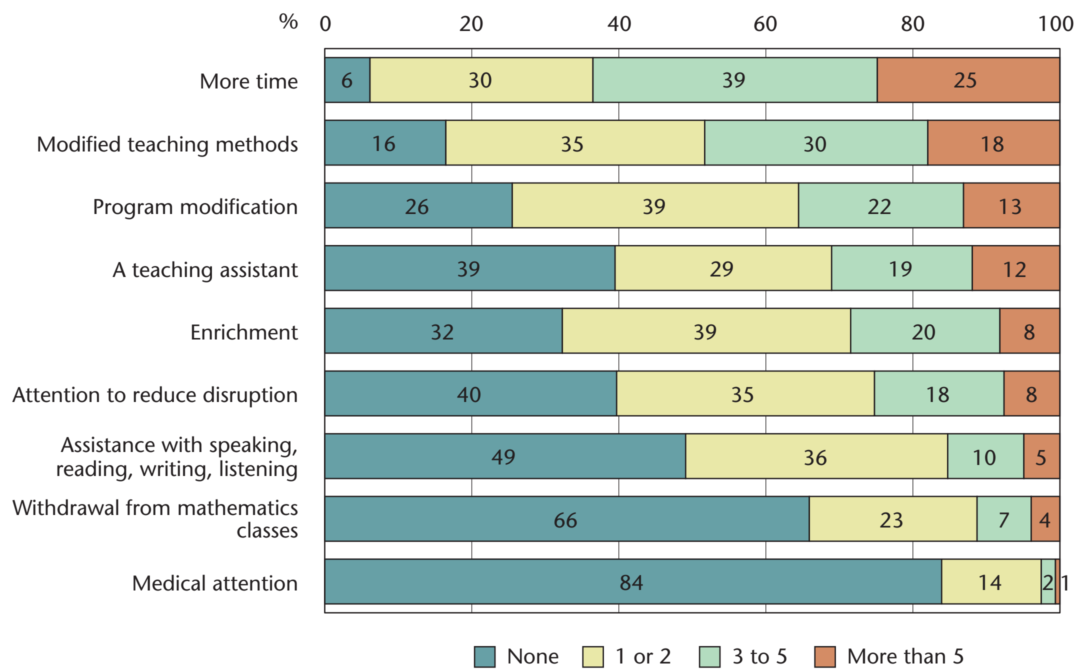
Chart 6-15 **Mean mathematics scores for special-needs placement and effects by jurisdiction and language**



The issue of accommodating special-needs students was examined in more detail in the teacher questionnaire. Teachers were asked a series of questions about the number of students in their mathematics classes who require various types of accommodation or intervention because of special needs. Chart 6-16 gives results for these questions for Canada. Results for populations vary widely but are quite complex and are thus not reported here.

The specific types of accommodations required are widely variable. More time, modified teaching methods and program modification are most common, while having students requiring withdrawal from class or medical attention are relatively rare.

Chart 6-16 **Teacher reports of number of mathematics students requiring accommodations for various special needs**



Teachers were also asked to what extent they adjust their teaching strategies to accommodate special-needs students and to what extent they see their classes affected by the presence of these students. Results for these two questions are given in Charts 6-17 and 6-18. On average, about two-thirds of teachers reported that they do not modify their strategies at all or only a little. The percentages responding "more than a little" or "a lot" varies substantially across populations.

As Chart 6-18 shows, a larger proportion of teachers indicated that the presence of special-needs students affects their classroom not at all or a little, again with substantial variation across populations. All jurisdictions reported a lesser affect of special-needs students in classrooms compared to that reported in the PCAP-13 2007 Contextual Report which may reflect the success of inclusion or differentiation initiatives.

Chart 6-17 **Extent of modification of mathematics teaching strategies for whole class to accommodate special-needs students by jurisdiction and language**

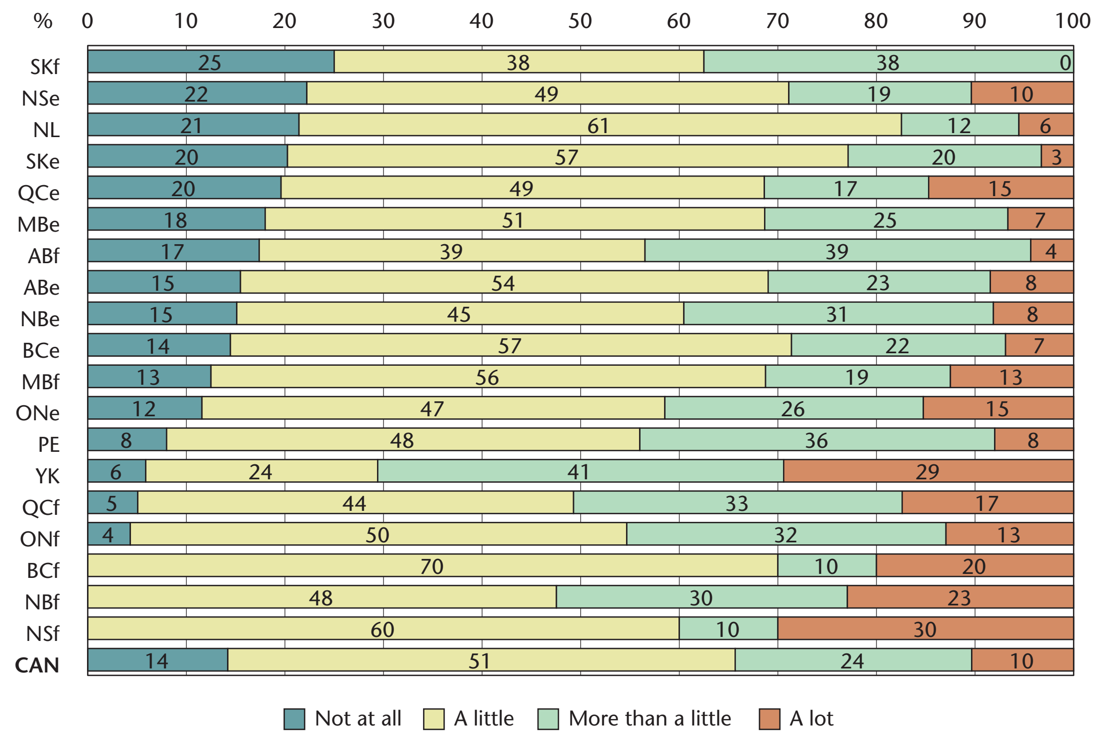


Chart 6-18 **Affects of special-needs students on mathematics classes by jurisdiction and language**

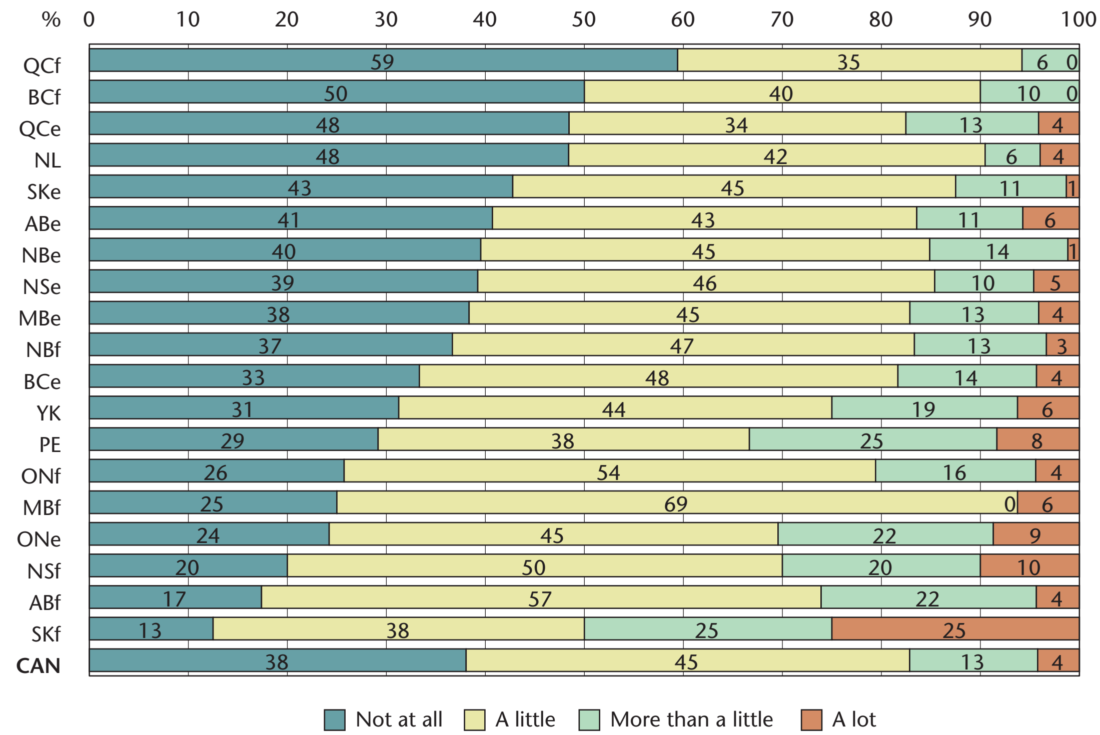
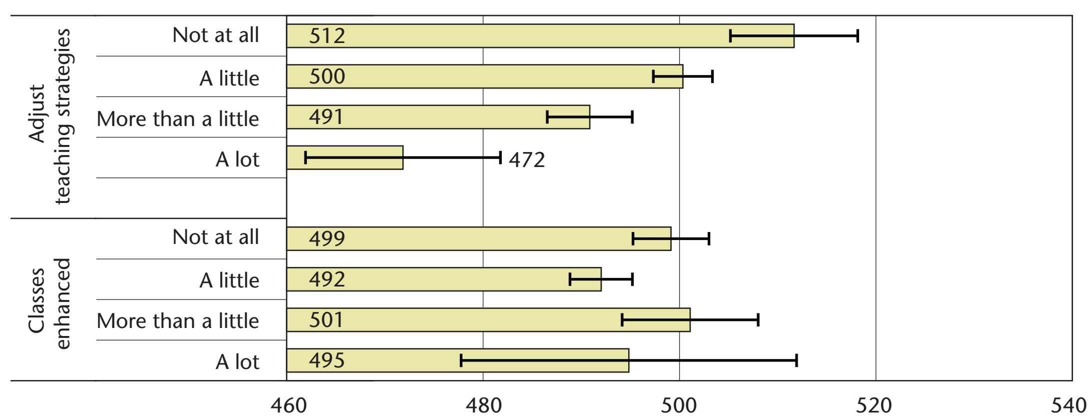


Chart 6-19 gives mean mathematics scores for adjusting teaching strategies and enhancement. Having to adjust teaching strategies has a significant negative effect on achievement across all categories, while enhancement has no effect.

Chart 6-19 **Mean mathematics scores for adjustment of teaching strategies and enhancement of classes by special-needs students**



The presence of an adult other than the regular teacher is characteristic of classes where special-needs students are included. Chart 6-20 gives the amount of time reported by teachers in which another adult is present in their mathematics classes to help individual students. Overall, for Canada, about half the teachers reported "none of the time." However, the overall distribution varies considerably by population.

Chart 6-21 shows that the presence of other adults in the classroom is negatively associated with achievement. The general pattern here is that of lower scores with increased time, although the difference between the two highest categories is not statistically significant.

CHART 6-20 **Presence in mathematics classes of an adult other than the regular classroom teacher by jurisdiction and language**

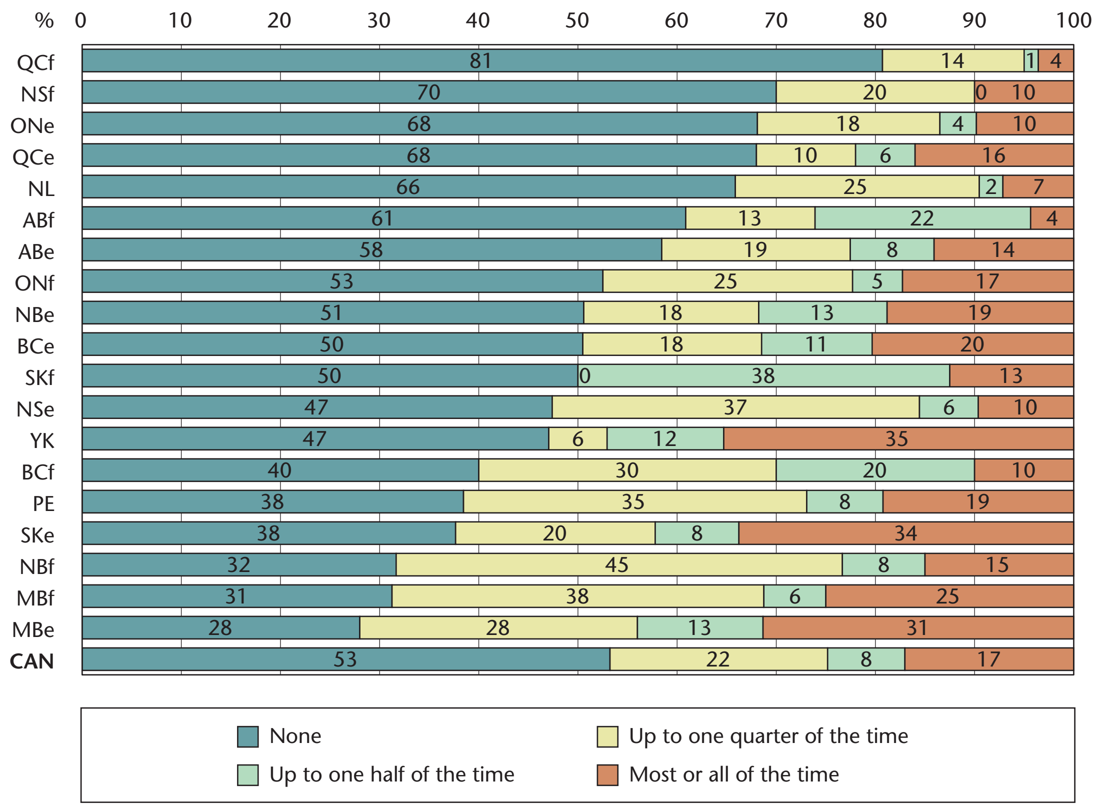
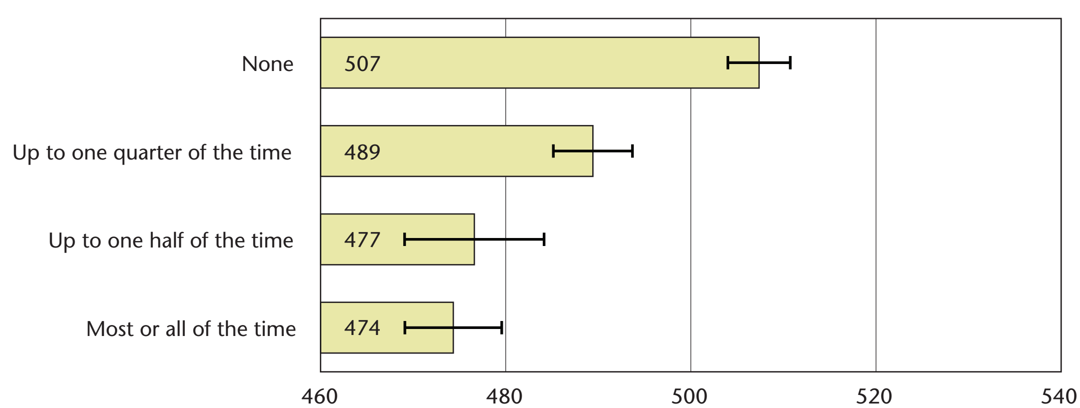


CHART 6-21 **Mean mathematics scores for presence in mathematics classes of an adult other than the regular classroom teacher**



Challenges in teaching mathematics

Teachers responded to a series of 20 questions on the challenges they face in teaching mathematics, using a three-point scale from "little or no challenge" to "a great challenge." These yielded six factors as described in Table 6-3.

TABLE 6-3 **Challenges in teaching mathematics factors and items**

Factors	Items
Resources	• Shortage of materials or equipment • Shortage of computer hardware or software • Inadequate physical facilities • Inadequate resource materials for lesson planning
Student background	• The range of student abilities in the class • Students coming from a wide variety of backgrounds • Students with special needs
Safety/morale (–)	• Concerns for personal safety or the safety of students • Low morale in the school
Teacher	• Limits in my own background in the subject • Lack of professional development • Lack of time for planning
Program (–)	• Too much content in curriculum • Curriculum inappropriate for grade level • External assessments or standardized tests • Weak curriculum • Large class sizes
Discipline (–)	• Disruptive students • Uninterested students • Pressure from parents/guardians
Note: (–) indicates factors that showed negative loadings of their items during the factor analysis.	

Charts 6-22 to 6-27 show the distribution of factor scores on these factors by population. These results may be summarized briefly as follows:

- For resource challenges, French populations in New Brunswick, Manitoba and Saskatchewan, as well as Ontario English, are above the Canadian average. Most of the Atlantic populations, as well as Alberta, English and French, and Saskatchewan English are below.
- Yukon stands out as above the Canadian average for student background challenges. A cluster of mainly francophone populations, along with Nova Scotia English and Newfoundland and Labrador, are below the Canadian average.
- Safety and morale challenges are greater than the Canadian average in the English populations in Nova Scotia, Ontario, Saskatchewan, and New Brunswick and in Ontario French. Most other francophone populations, as well as Manitoba English, Quebec English, and British Columbia English, are below.

- Four of the francophone populations, British Columbia, Saskatchewan, New Brunswick, and Ontario, along with Yukon and Saskatchewan English, are above the Canadian average on teacher challenges. English populations in Newfoundland and Labrador, Nova Scotia, Quebec, British Columbia, and New Brunswick are below the Canadian average.
- Program challenges show wide variations, with seven populations above and seven below the Canadian average.
- Finally, discipline challenges show a clear cluster of four francophone jurisdictions, along with Prince Edward Island, with higher scores than the Canadian average and higher than any others. Most others are at the Canadian average, with only three — New Brunswick French, Nova Scotia English, and Quebec French — below.

Chart 6-22 **Resource challenges in teaching mathematics by jurisdiction and language**

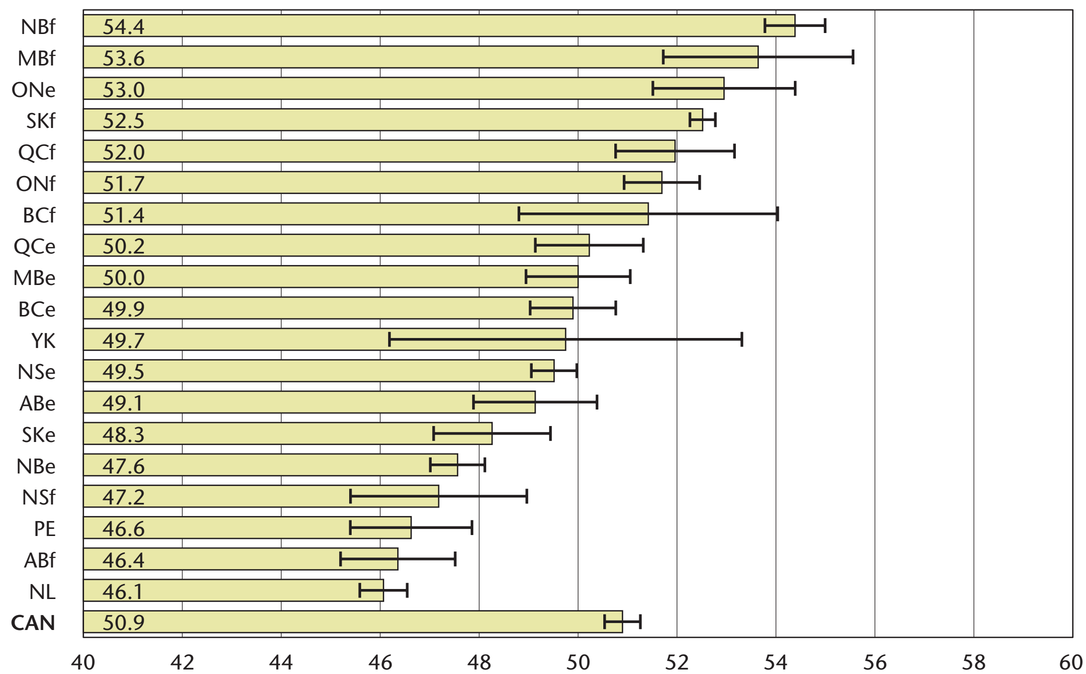


Chart 6-23 **Student background challenges in teaching mathematics by jurisdiction and language**

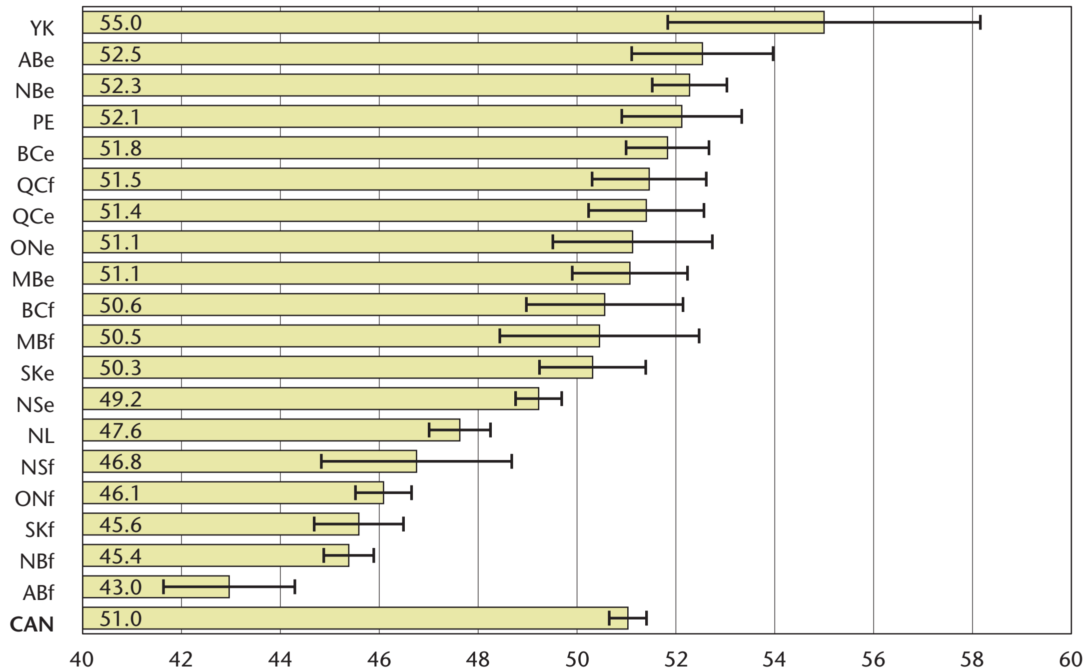


Chart 6-24 **Safety/morale challenges in teaching mathematics by jurisdiction and language**

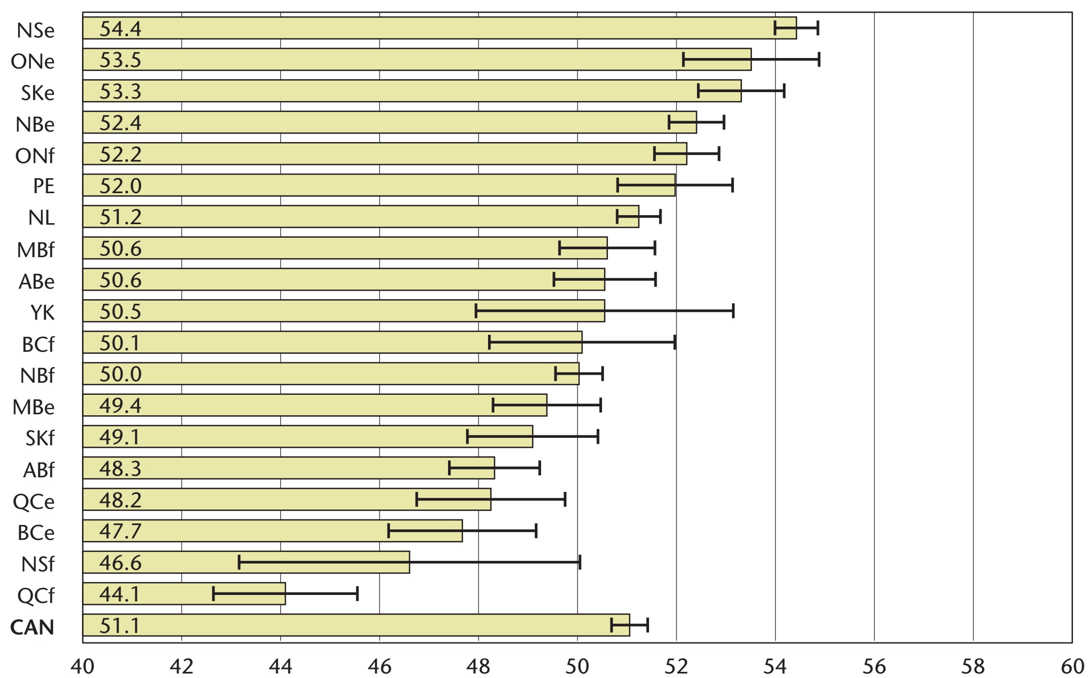


CHART 6-25 **Teacher challenges in teaching mathematics by jurisdiction and language**

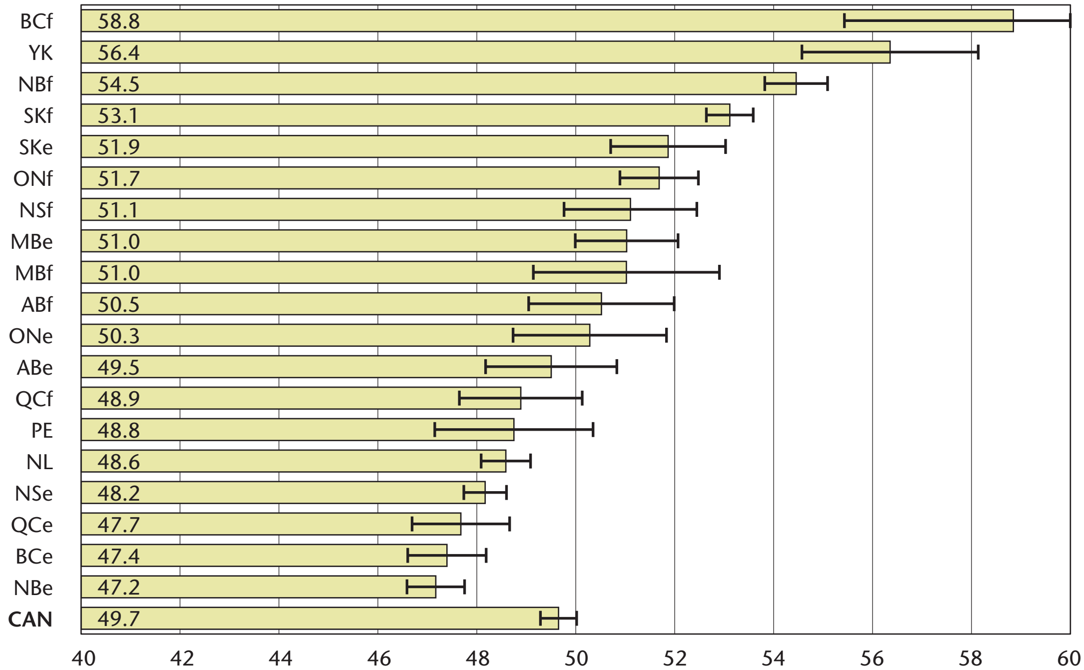


CHART 6-26 **Program challenges in teaching mathematics by jurisdiction and language**

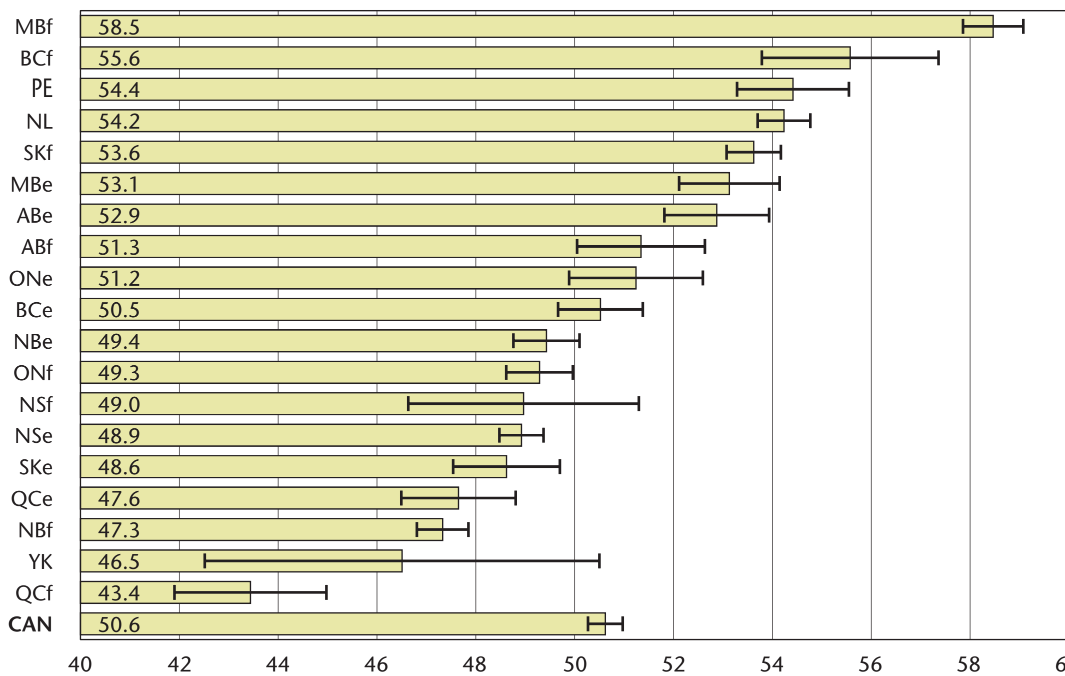
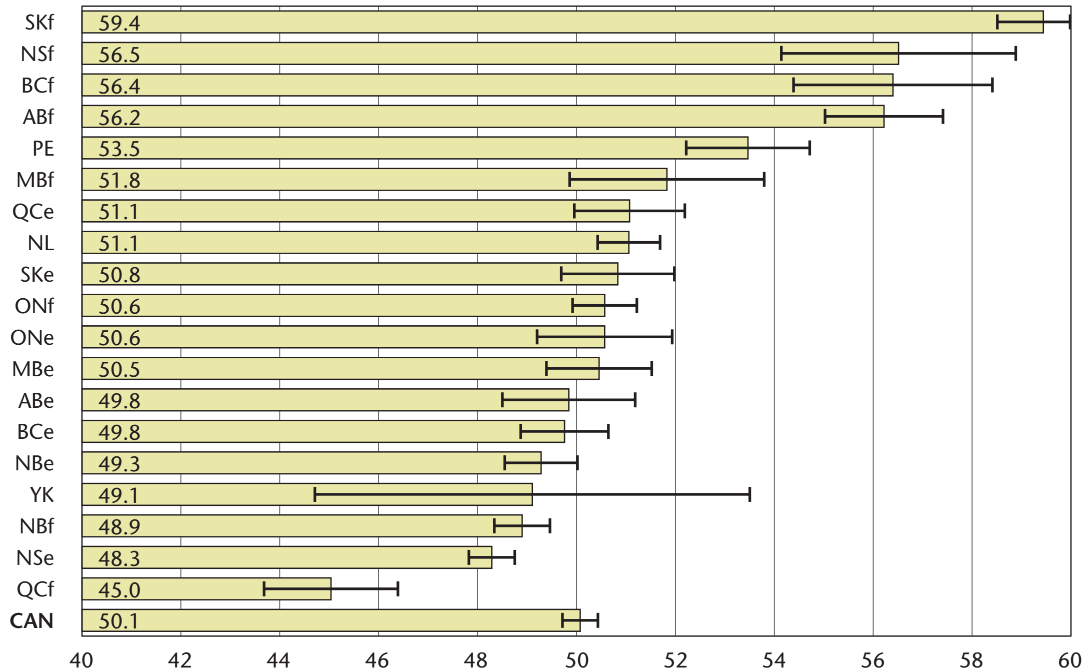
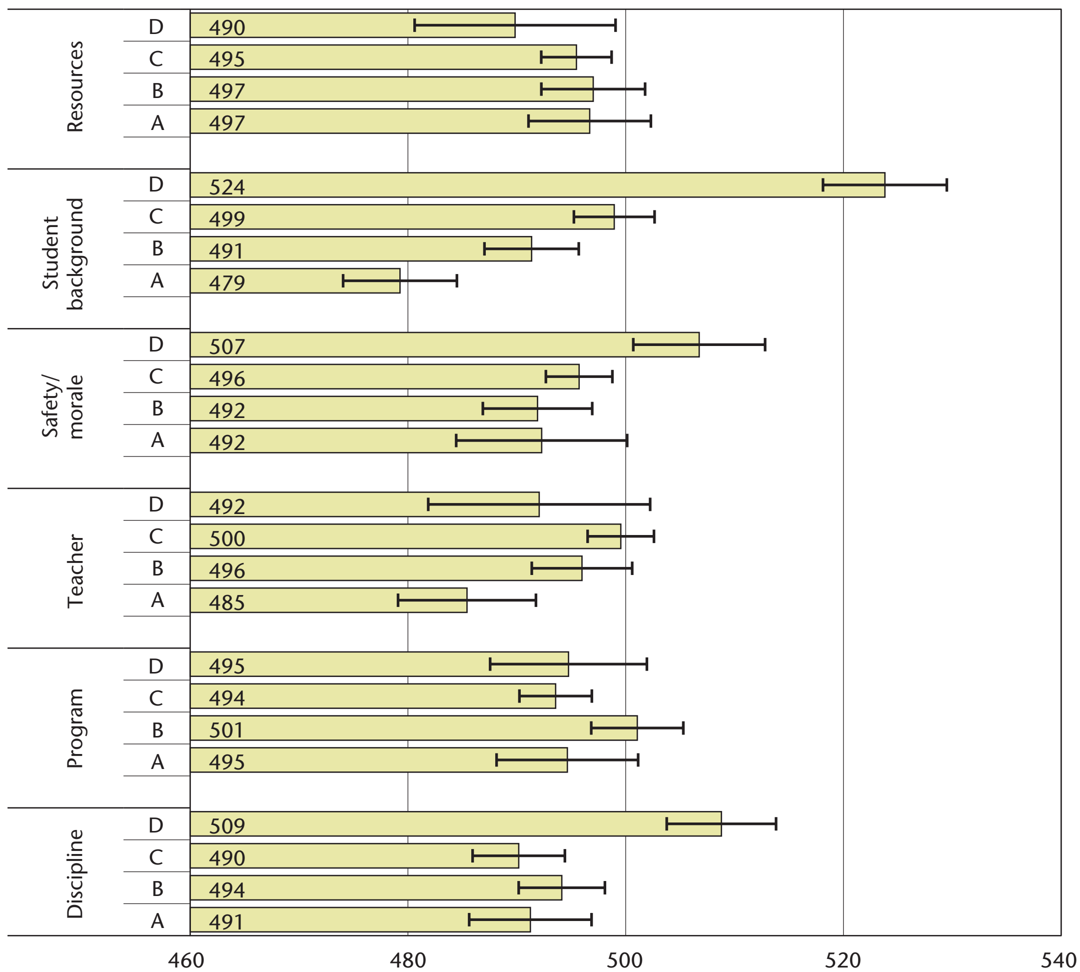


Chart 6-27 **Discipline challenges in teaching mathematics by jurisdiction and language**



Changes in mean mathematics scores for one standard deviation change in these challenges are given in Chart 6-28. The most distinct pattern is for student background challenges, which shows a strong negative effect. In other words, when teachers perceived that student background posed little or no challenge to teaching, higher mean math scores were found. Safety/morale and discipline challenges also show negative effects, statistically significant only for the lowest level on these factors, relative to all other levels. For the teacher factor, a significant difference is found between levels C and A, with higher scores for the lower level (level C). Resource and program challenges show no significant effects. These results point to a broader challenge for policy in that the challenges showing significant effects are those least amenable to change through policy or programming as they more likely originate outside the school.

Chart 6-28 **Mean mathematics scores for challenges in teaching mathematics**



Disciplinary climate

Students were asked the following three questions about lost time and disruption in mathematics classes:

How often do the following occur in your mathematics classes?

– We lose time because of student misbehaviour.

– We lose time because of other disruptions (e.g., announcements, visits).

– We lose time because of discussions unrelated to the mathematics lesson.

These yielded a single factor, labelled "disciplinary climate." High scores on this factor correspond to more time lost and greater disruption and thus to a poorer disciplinary climate.

Chart 6-29 shows mean factor scores by population. This shows two populations, Nova Scotia French and Quebec French, at the top of the scale, significantly higher than the Canadian average and significantly higher than any others. Nova Scotia English is also above the Canadian average. Four populations — New Brunswick French, British Columbia English, Manitoba French, and Saskatchewan French — are below the Canadian average and also lower than any others.

Chart 6-29 **Mean factor scores for disciplinary climate in mathematics classrooms by jurisdiction and language**

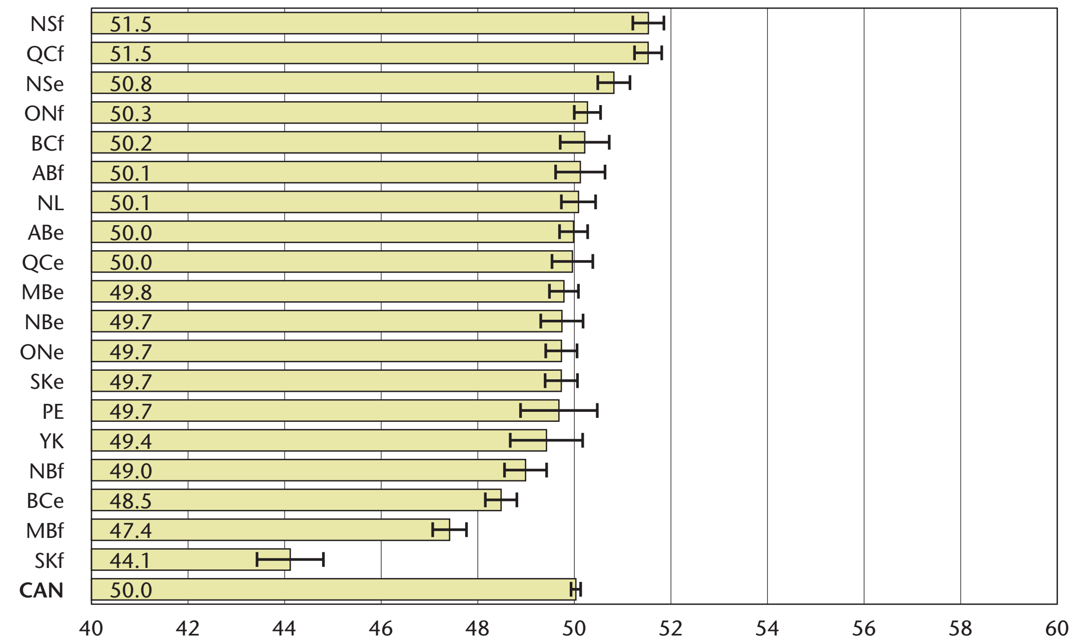
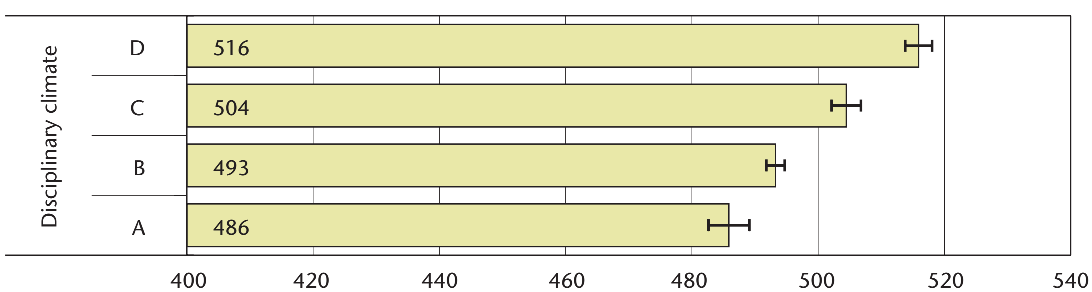


Chart 6-30 gives the mean mathematics scores by standard deviation unit for disciplinary climate. There is a clear linear pattern here in which a more negative disciplinary climate is associated with lower scores.

Chart 6-30 **Mean mathematics scores for disciplinary climate in mathematics classrooms**



Multiple regression effects

The effects of the various blocks of variables used in this chapter were examined, as before, using multilevel modelling. In this case, the "areas of emphasis" cluster and the "sources of influence" cluster were dropped from the model because each of these accounted for less than 2 per cent of either the student-level or the school-level variance. Disciplinary climate and class size also were not modelled because these variables are not part of any of the other clusters and thus should not be part of a model designed to investigate the relative effects of variables within clusters. Multiple regression effects for these variables will appear in the "full" model presented in the final chapter.

Chart 6-31 gives the coefficients for the variables related to how special-needs students are accommodated. Significant effects for both models are found for placement of special-needs students with the regular teacher, the extent of modification of teaching to accommodate special needs, and the presence of an adult other than the regular teacher in the classroom. Also, there are no statistically significant changes from the simple to the multiple regression model, indicating that these variables are independent of each other in their effects on mathematics scores.

CHART 6-31 **Regression coefficients for accommodation of special-needs students**

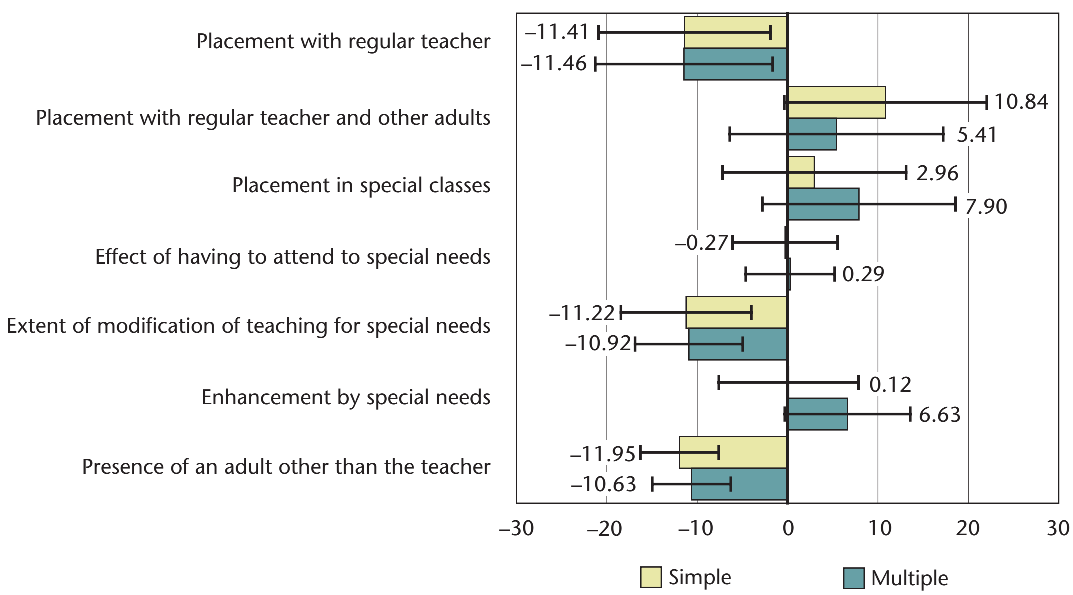
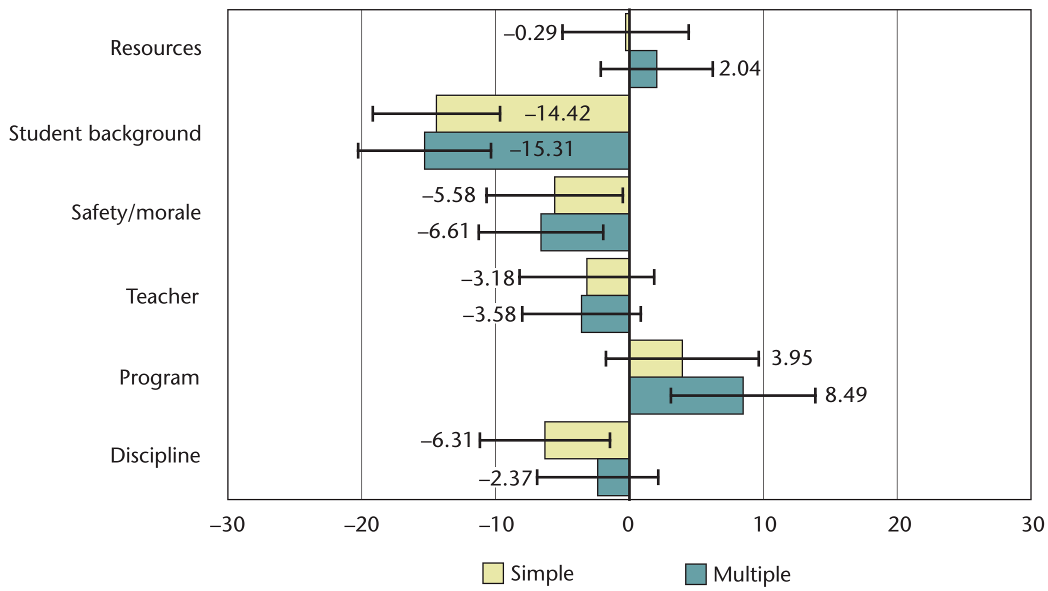


Chart 6-32 shows the effects on mathematics scores of challenges to mathematics teaching. In this case the coefficients are directly comparable because all variables are on the same scale. Student background again shows the strongest negative effect in both models. The effect for safety/morale challenges also remains negative in the multiple regression model. On the other hand, the program effect becomes significantly positive when the other challenges are controlled. Discipline shows the opposite pattern, going from significantly negative in the simple regression model to non-significant in the multiple regression model. In neither of these cases, however, is the change from the simple to the multiple regression model statistically significant. Considering the width of the confidence intervals, these effects should therefore be treated as marginal.

Chart 6-32 **Regression coefficients for challenges in mathematics teaching**



7 TIME ALLOCATION AND USE

All learning may be thought of as occurring within a time framework. At the broadest policy level, the length of school years and of school days is established through legislation. Time spent on subjects is also sometimes determined provincially/territorially. At the school and classroom levels, many activities are part of the schedule, and trade-offs are often necessary because total time is fixed. Individual students may spend more or less time on school work, both within the classroom (engagement) and outside (homework or other school-related activities). While not all of these time elements can be captured in a broad survey, questions on many aspects of time are found in all of the PCAP questionnaires.

School time on mathematics

Principals were asked to estimate the number of minutes per week spent on mathematics in their school. Estimates were wide-ranging but tended to cluster around several modal points such as 200 or 300 minutes. For ease of presentation, the estimates were divided into four categories. The results by jurisdiction and language are given in Chart 7-1. This shows the wide variation among populations and also wide variation across schools within most populations. New Brunswick French has the most schools in the highest range of more than 300 minutes, followed by Yukon and Prince Edward Island. British Columbia English stands out as having many schools at both extremes of the distribution and few in the mid-ranges. On the other hand, several populations, from Ontario English to Alberta English on the graph, have few schools at either extreme and most in the mid-ranges.

Mean mathematics scores by variations in mathematics time are given in Chart 7-2. Although there is no strong pattern here, schools in the range 201–250 minutes per week have significantly higher scores than those who spend more time than this on mathematics. While this seems counterintuitive, it is possible that time on mathematics is confounded with other variables. For example, it is possible that schools with lower achieving students assign more time to core subjects. The multiple regression models may shed more light on this issue.

Chart 7-1 **Minutes per week on mathematics instruction by jurisdiction and language**

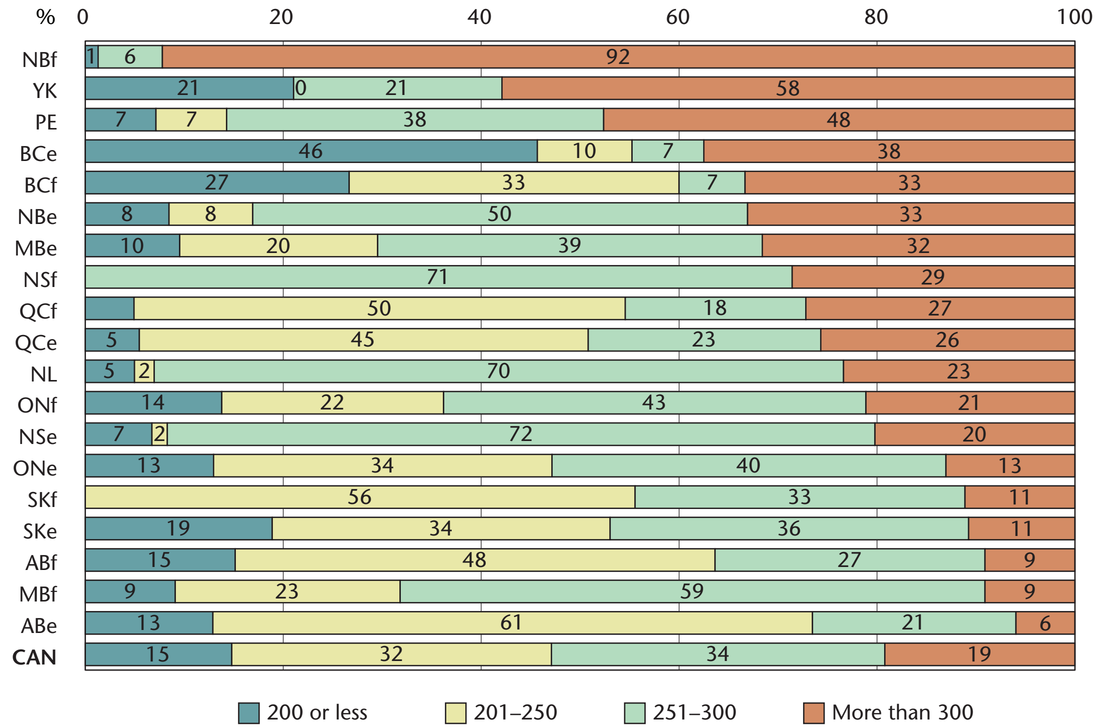
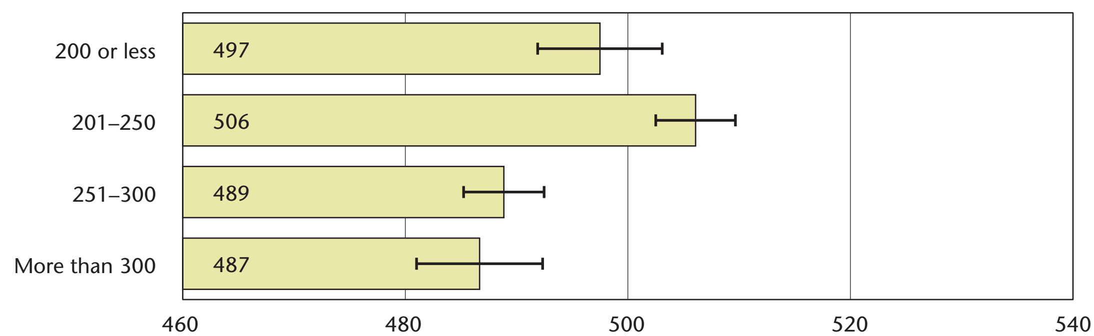


Chart 7-2 **Mean mathematics scores by minutes per week on mathematics**



Length of class periods is sometimes considered to be a useful indicator of efficiency of time use, because longer periods result in less transition time. Again, this shows considerable variation both within and across populations, as indicated in Chart 7-3. In most populations, relatively few schools are at either extreme. Notable exceptions are Manitoba English, with close to half the schools having class periods 40 minutes or less, and Yukon and British Columbia English with relatively large numbers of schools with more than 75 minutes.

Chart 7-4 gives the mean mathematics scores by length of class period. Again, there is no clear trend except that schools with more than 75-minute periods have lower scores than those having any shorter length.

Chart 7-3 **Average minutes in class periods by jurisdiction and language**

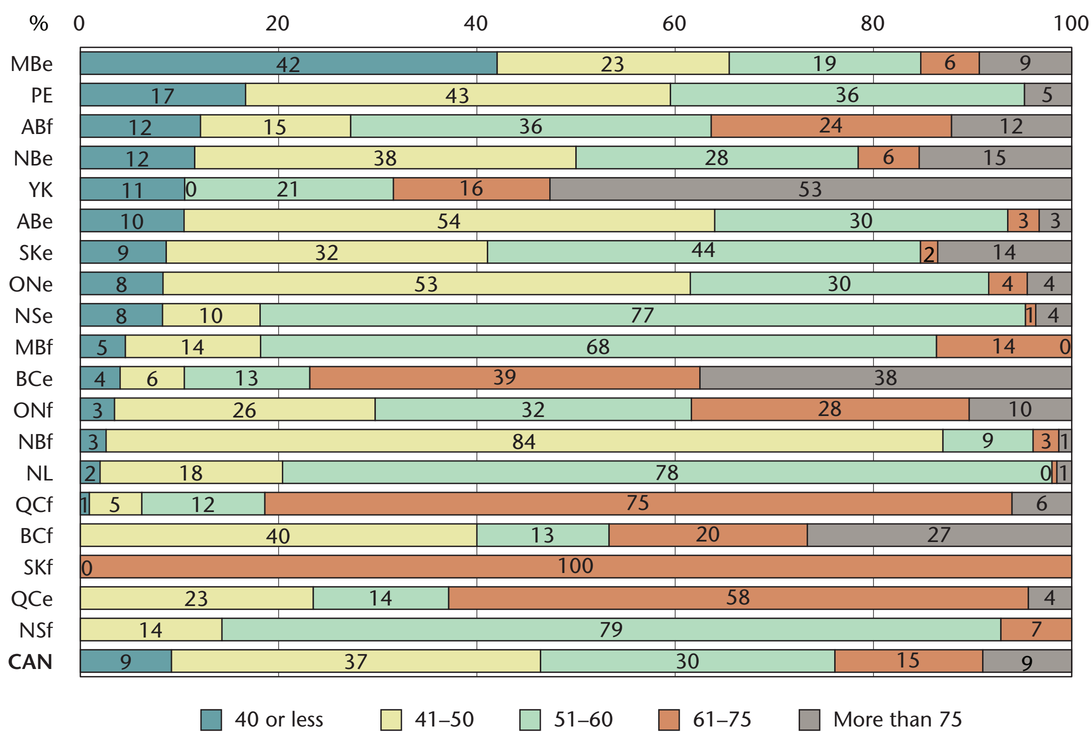
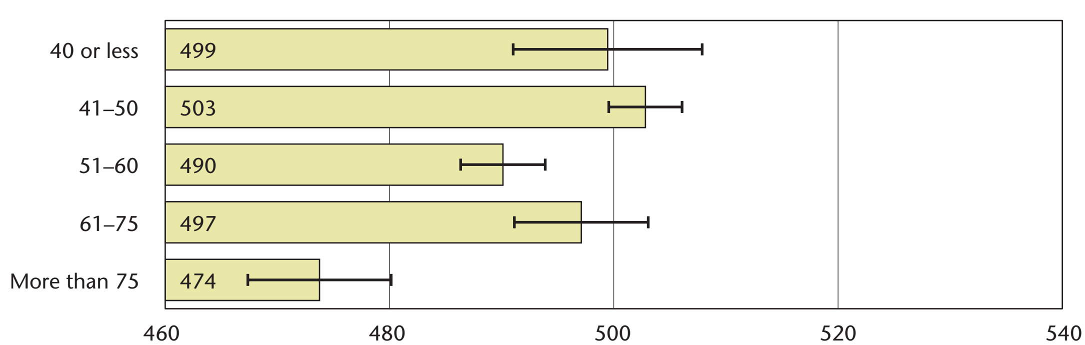


Chart 7-4 **Mean mathematics scores by minutes in class periods**



Student absence

Data on student absence were available from both the school and student questionnaires. School absence rates by jurisdiction and language are shown in Chart 7-5. There is a distinct clustering by language in this case, with almost all schools in francophone populations reporting an average absence rate of less than 5 per cent . Yukon stands out at the opposite end, with more schools than in other jurisdictions reporting a more than 10 per cent absence rate. (Again, it must be noted that the total number of schools in Yukon is very small.)

The relationship of school absence rates to school mean mathematics scores is shown in Chart 7-6. A clear linear pattern is evident, with schools having higher absence rates having significantly lower scores.

CHART 7-5 **School absence rates by jurisdiction and language**

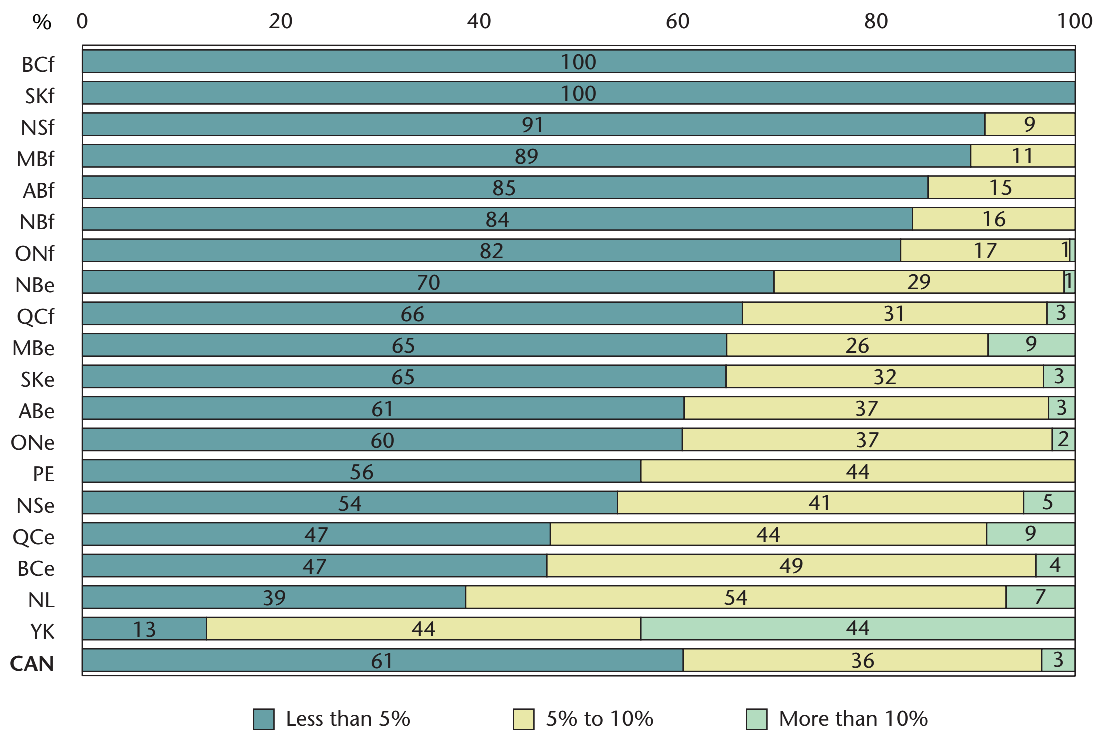
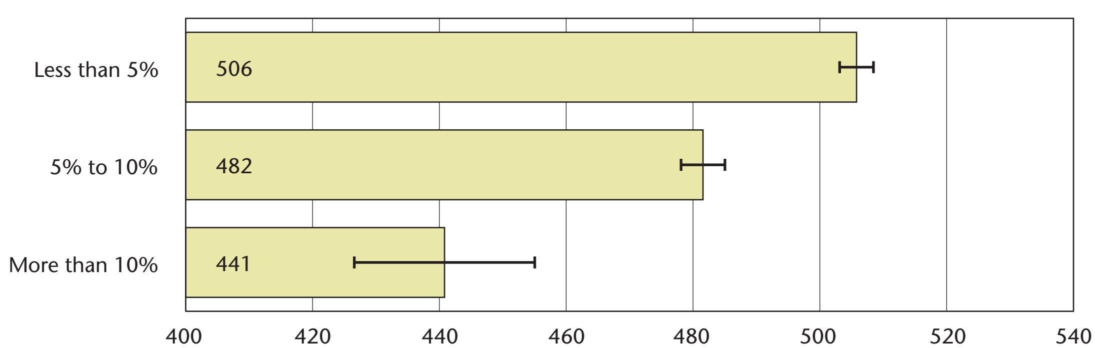


CHART 7-6 **School mean mathematics scores by school absence rates**



Students were asked to report the number of days they had been absent during the current school year for both non–school-related and school-related (field trips, sport events, etc.) reasons. Breakdowns by jurisdiction and language are given in Charts 7-7 and 7-8. For non–school-related absences, the distribution is fairly even across the four categories used and is not very different across populations. Absences for school-related reasons are less frequent generally but are more variable across populations.

Chart 7-7 **Student absence for non–school-related reasons by jurisdiction and language**

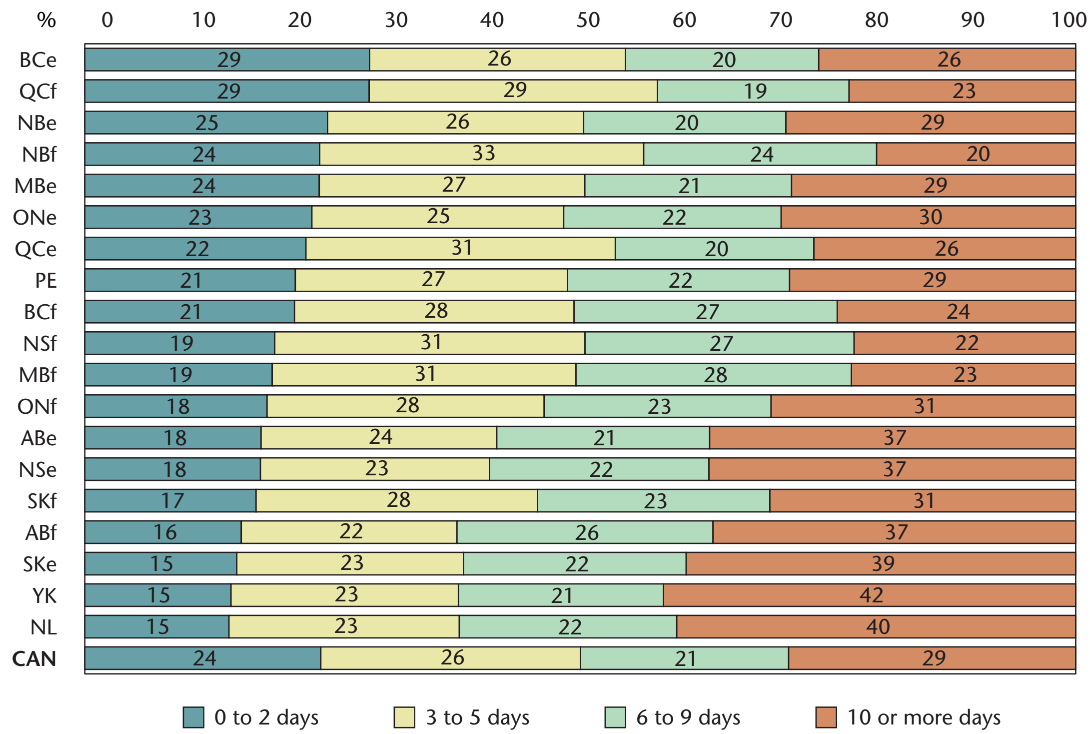


Chart 7-8 **Student absence for school-related reasons by jurisdiction and language**

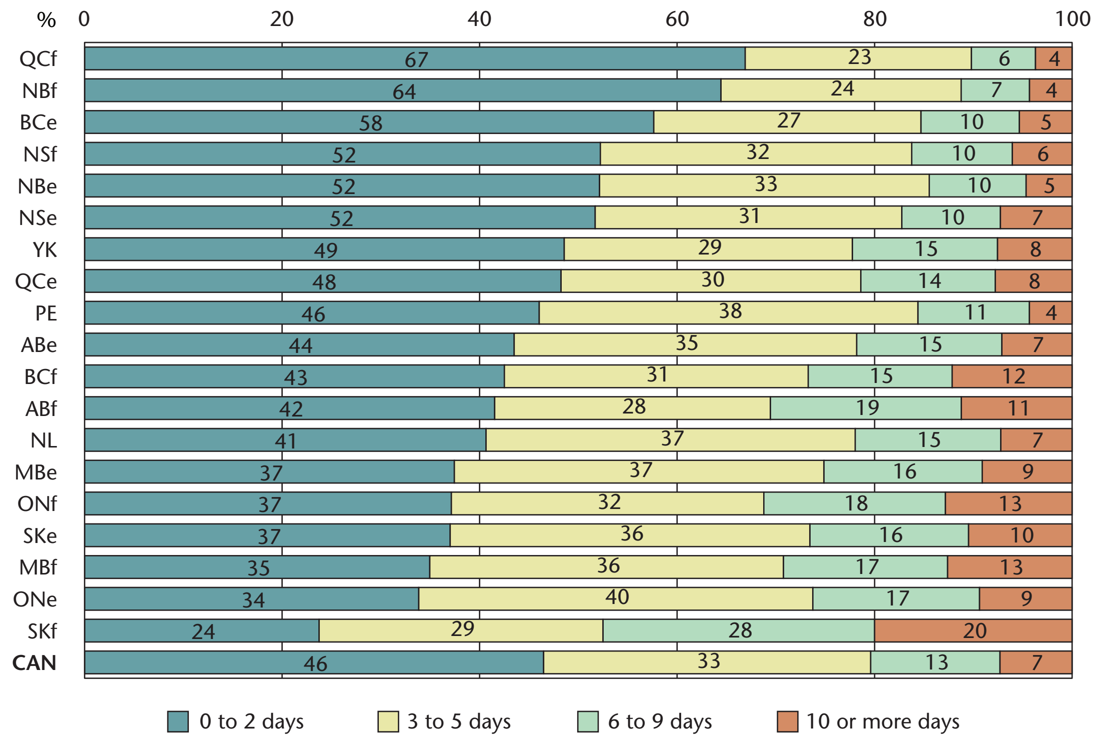
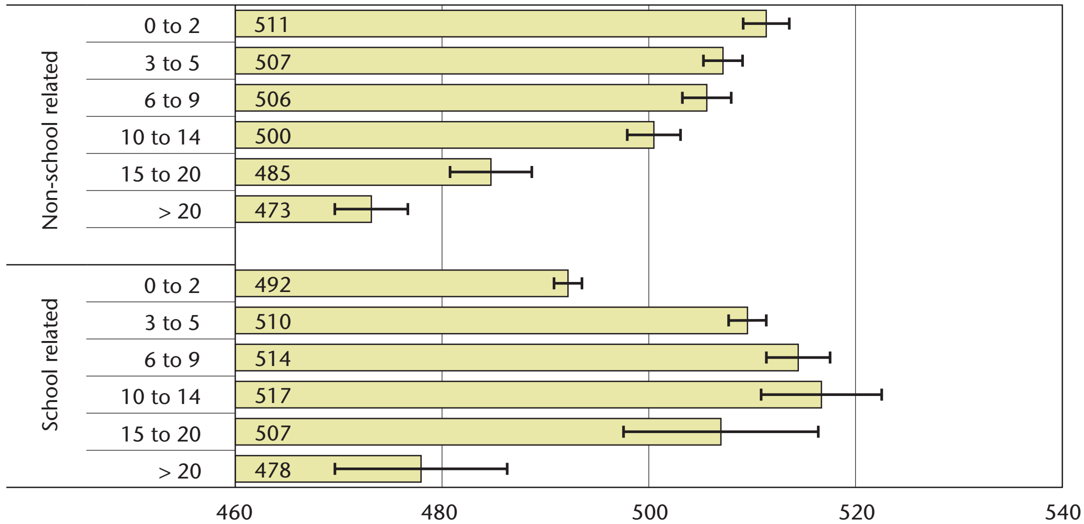


Chart 7-9 gives mean mathematics scores for students at various numbers of days absent. In this case, a wider range of categories is used to examine more closely what happens to those few students who have a high rate of absence. Absence for non–school-related reasons shows a distinct pattern of lower scores for higher absence rates. The trend accelerates for those in the two highest absence categories. Even though relatively small numbers of students are in these categories, they have significantly lower scores than students with lower absence levels.

The pattern for school-related absence is non-linear, with the lowest scores being achieved by those in the two extreme categories. This raises the interesting issue that some absence due to involvement in field trips, sports, music, and the like is desirable, but that this becomes counterproductive if it reaches extreme levels.

Chart 7-9 **Mean mathematics scores by student absence**



Homework

Information on homework was gathered from both teachers and students. Teachers were asked about how much time they expect students to spend on homework and the types of homework they assign. Students were asked to report the amount of time they spend on homework in total and in mathematics.

Teacher reports of expected minutes per week of mathematics homework are given in Chart 7-10. Almost all teachers in all populations expect at least some mathematics homework, with the largest proportions in most cases expecting either 30–60 minutes or 60–120 minutes. Variations across populations are relatively large for the second category (less than 30 minutes), but are smaller for the other categories.

CHART 7-10 **Teacher-expected minutes per week mathematics homework by jurisdiction and language**

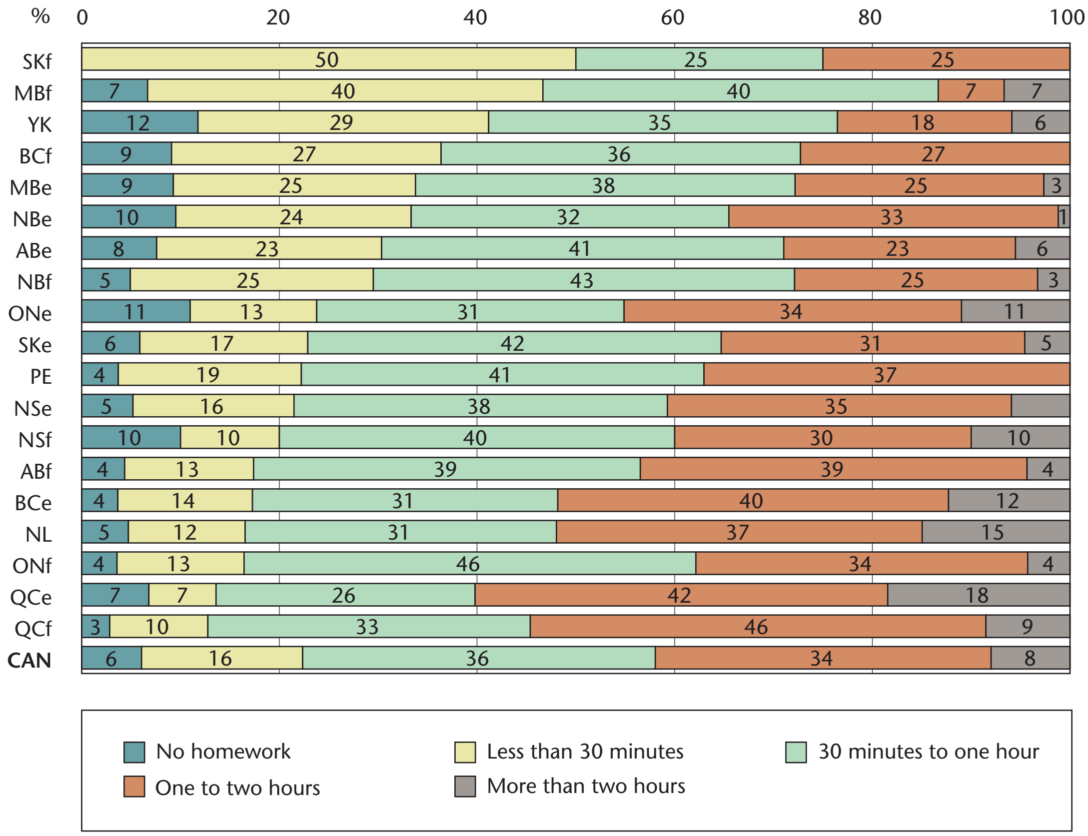
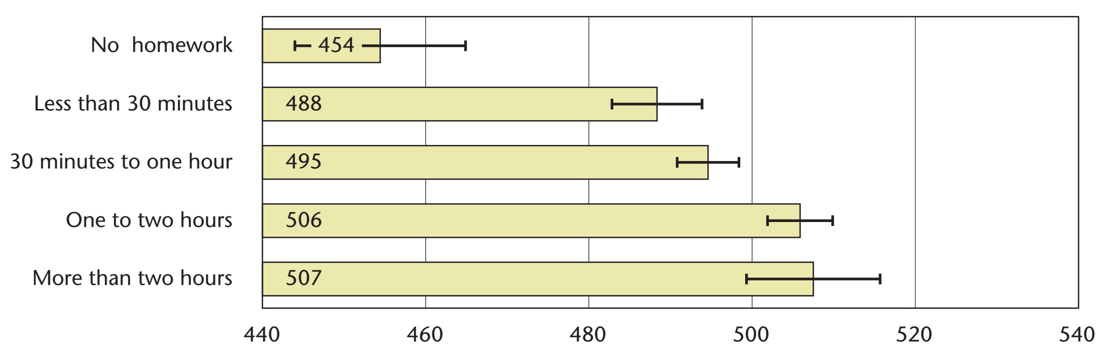


Chart 7-11 gives mean mathematics scores for teachers reporting the various homework ranges. The general pattern is that of higher mathematics performance in classes where more homework time is expected. The largest differences are between those having no homework and all others. The results also suggest a saturation point, with more than two hours of homework not yielding any further gain in performance.

CHART 7-11 **Mean mathematics scores by teacher-expected minutes per week mathematics homework**



Teachers were also asked a number of questions about the type of homework assignments given in mathematics, on a four-point scale from "rarely or never" to "almost every class." These questions yielded a three-factor pattern as shown in Table 7-1. Mean factor scores for each of these factors by population are given in Charts 7-12 to 7-14.

The drill and practice factor shows fairly wide variation with seven populations above the Canadian average and six below. The projects factor shows even greater variation but a more distinct clustering with Saskatchewan French, British Columbia French, and Ontario English being above the Canadian average and above most other populations, and Quebec English, Yukon, and especially Quebec French standing out as below the Canadian average and below all others. Finally, the test preparation factor shows two clusters near the top, with Yukon, Alberta French, and Nova Scotia French being above all others. Saskatchewan French and British Columbia English and French form a second cluster, below the top three but above the Canadian average. Most of the remaining populations are close to the Canadian average, with Ontario French and Prince Edward Island being significantly below.

Table 7-1 **Types of mathematics homework assignments**

Factors	How often do you assign the following types of homework?
Drill and practice	• Practice • Problems to solve • Drill
Projects	• Creating problems • Projects • Activities using manipulatives
Test preparation	• Studying for tests • Practice tests or quizzes

Chart 7-12 **Drill and practice as mathematics homework by jurisdiction and language**

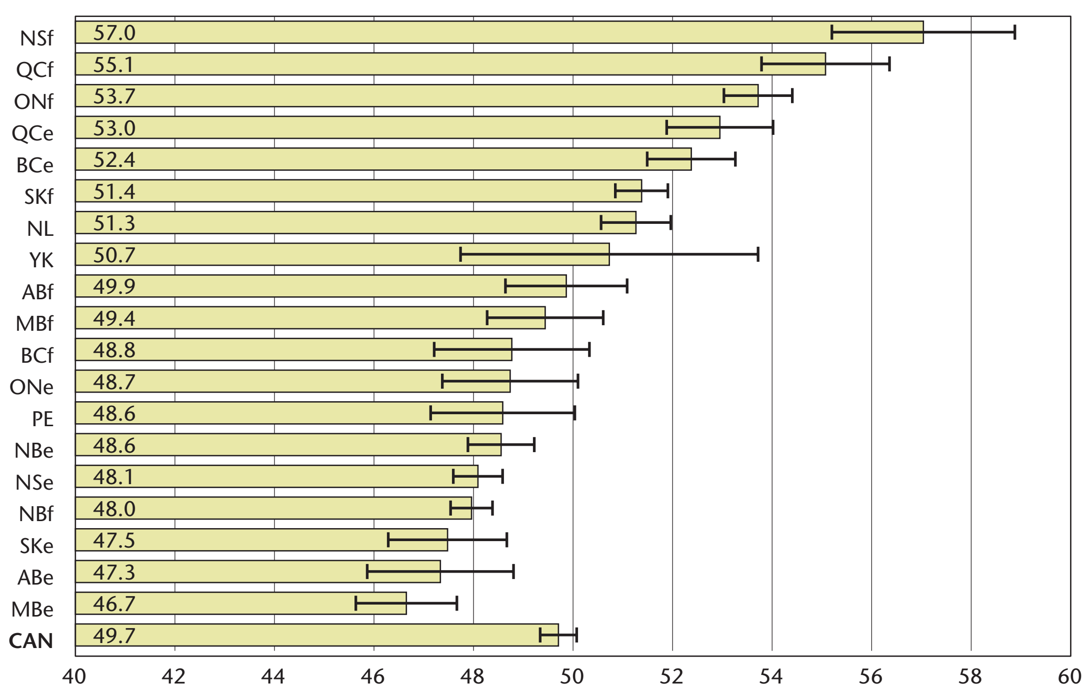


Chart 7-13 **Projects as mathematics homework by jurisdiction and language**

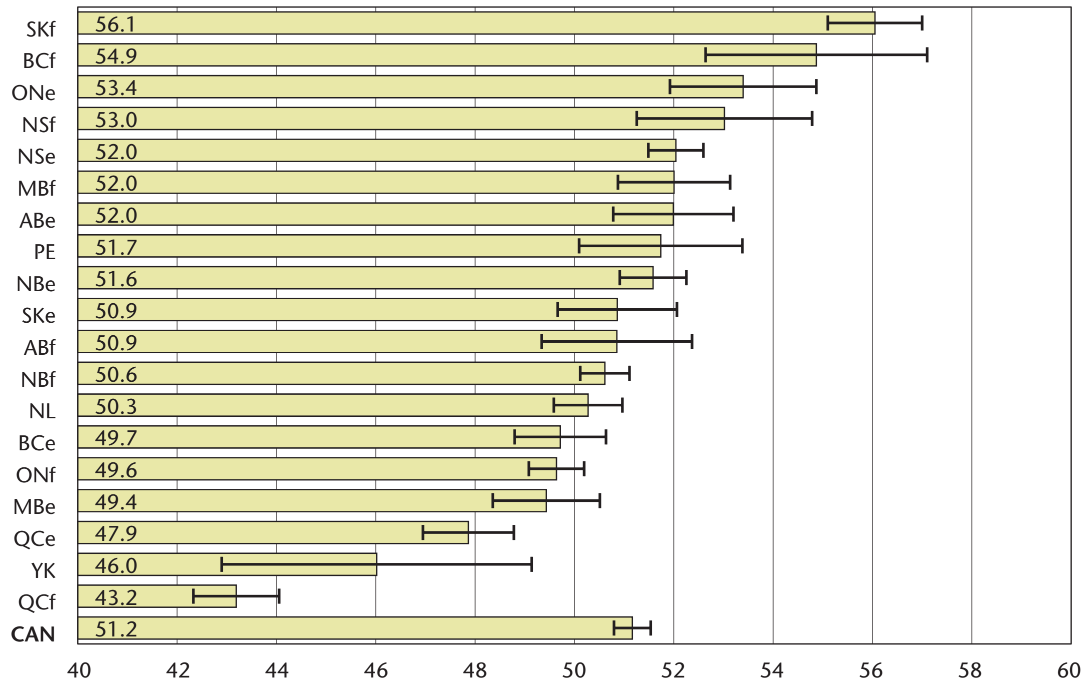


Chart 7-14 **Test preparation as mathematics homework by jurisdiction and language**

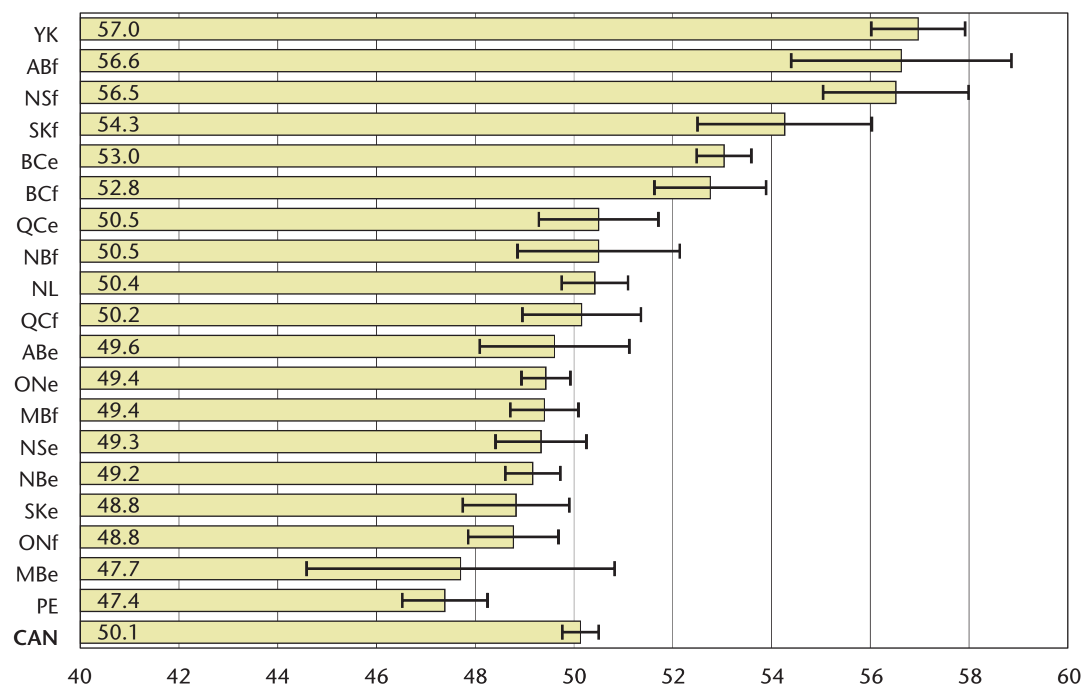
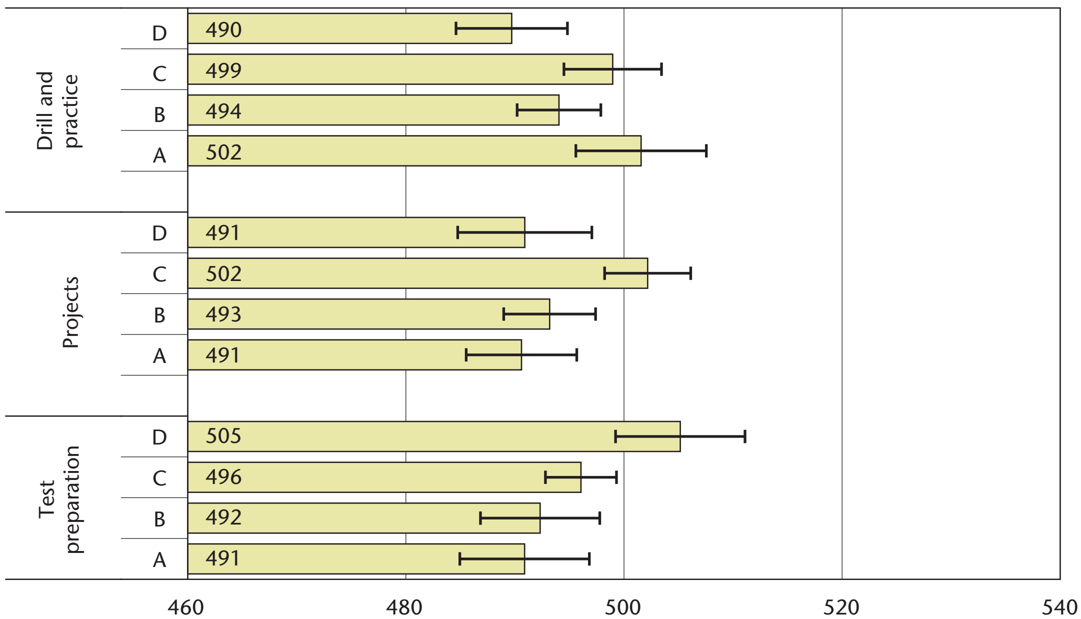


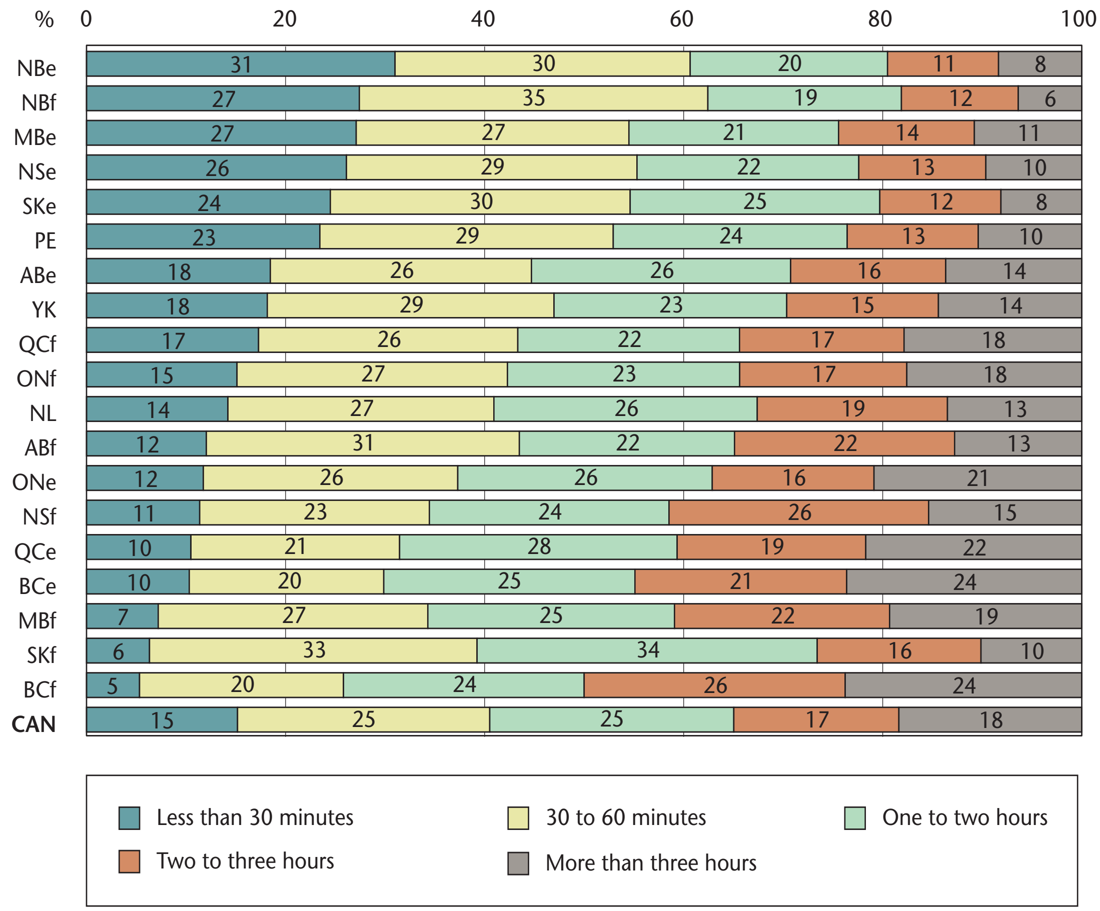
Chart 7-15 gives mean mathematics scores for homework types. Drill and practice shows no distinct pattern. Project work shows a non-linear pattern, with those in category C (the range −1 to 0 standard deviations) having significantly higher scores than those in other categories. For test preparation, the pattern is generally negative, with category D (those with the least amount of this type of homework) performing significantly better than categories B and A.

Chart 7-15 **Mean mathematics scores by types of mathematics homework**



Charts 7-16 and 7-17 present student reports of total minutes per week spent on all homework and on mathematics homework. There is wide variation both within and between populations in the proportions of students reporting various amounts of total homework. There is somewhat less variation across categories and across populations for mathematics homework, with a majority in all populations reporting 60 minutes or less on mathematics homework per week.[18] Relatively few students in any of the populations spend more than two hours per week on mathematics homework.

Chart 7-16 **Student weekly total homework time by jurisdiction and language**



[18] Total homework and mathematics homework are, of course, highly correlated because total homework includes mathematics homework.

Chart 7-17 **Student weekly mathematics homework time by jurisdiction and language**

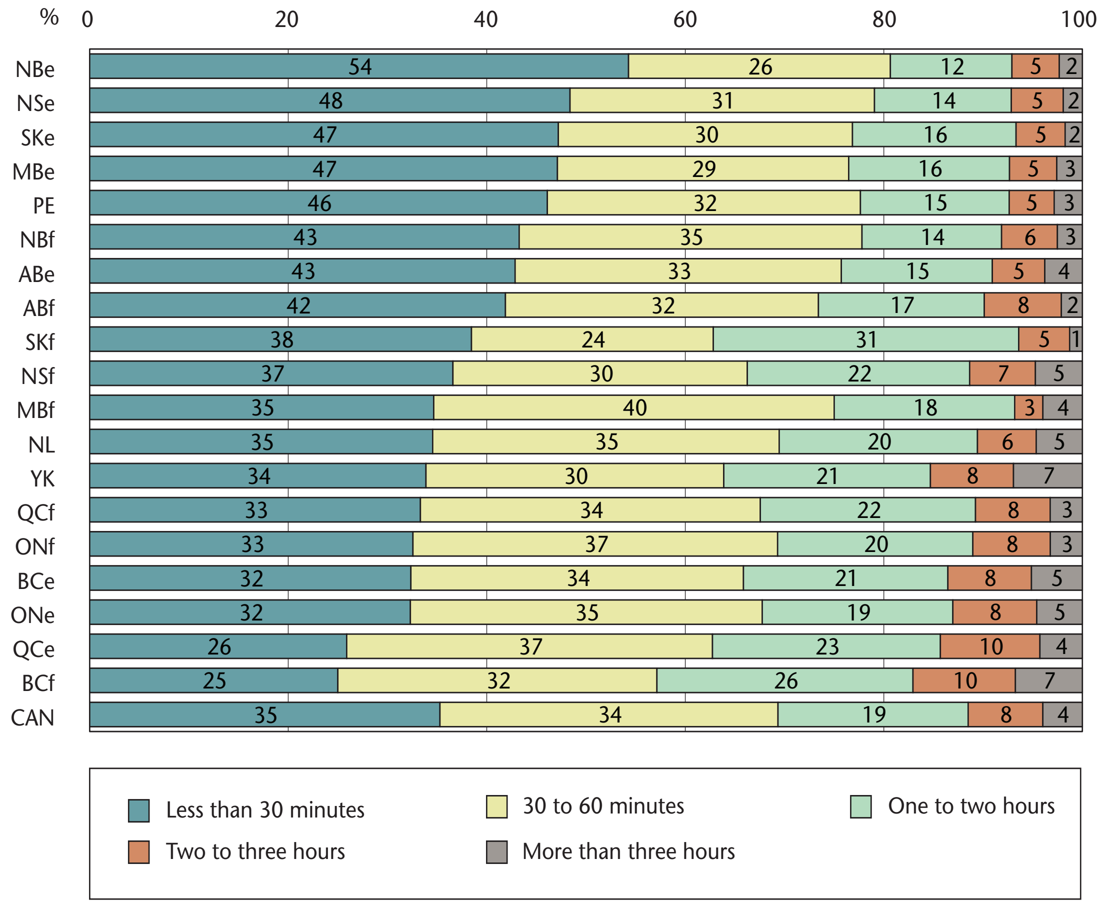
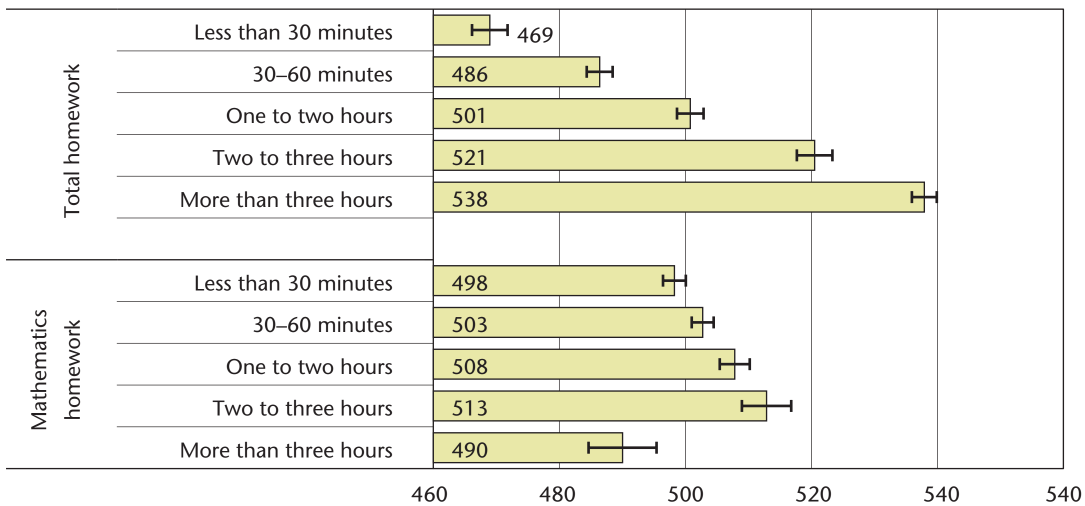


Chart 7-18 gives mean mathematics scores for students at different weekly amounts of homework. Total homework shows a strong trend toward higher scores for more time spent on homework. The pattern is in the same direction but is less pronounced for mathematics homework. This pattern is similar to that for teacher homework expectations, except that doing too much mathematics homework is counterproductive, with the score pattern being reversed at more than three hours. It is possible, of course, that students who are struggling with mathematics are the ones doing the most homework. If so, the results indicate that doing a large amount of homework is insufficient to transform low performance into high performance, even if the added homework does have some positive effect on individuals.

Chart 7-18 **Mean mathematics scores by weekly homework amounts**



Multiple regression effects

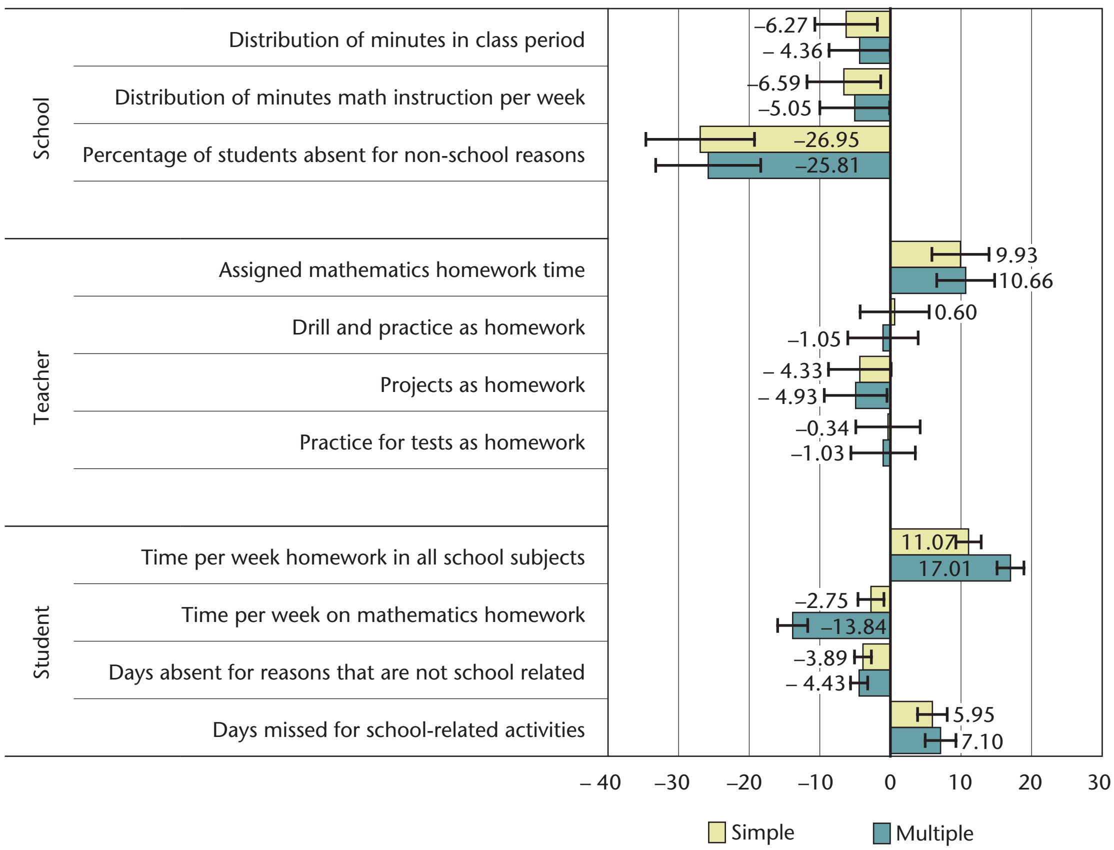
Chart 7-19 shows the simple and multiple regression effects[19] for the time allocation and use variables.[20] Most of these variables are statistically significant in both models and show no significant shift across the two models, indicating that these effects are largely independent of one another. The percentage of students absent for non-school reasons shows a relatively strong negative effect in both models.

For the two homework time variables, both the positive effect of total homework and the negative effect of mathematics homework increase from the simple to the multiple regression model. This reflects the non-independence of these variables and particularly the fact that mathematics homework is included in total homework. Thus, once total homework is controlled, mathematics homework would be expected to have no effect or a negative effect. The small but significant negative coefficient for mathematics homework in the simple regression model relates to the non-linearity of the simple regression effect, as earlier shown in Chart 7-18. Here the reduced score for those reporting more than three hours per week of mathematics homework, is sufficient to offset the smaller positive trend for the other homework time categories.

[19] Please refer to the Statistical Note in Chapter 3 for an explanation of the simple and multiple regression models.

[20] It is reiterated that the coefficients for different variables cannot be compared to each other because they are not all on the same scale. The simple and multiple regression models for the same variable are directly comparable.

Chart 7-19 **Regression coefficients for time allocation and use**



8 TEACHING AND LEARNING STRATEGIES IN MATHEMATICS

Mathematics teaching strategies

Data on mathematics teaching strategies were gathered from teachers using 10 items on a four-point qualitative time scale from "not at all" to "a lot." Factor analysis of these items gave three broad strategies as described in Table 8-1. The first two essentially distinguish between teachers who instruct through group work with students sharing problems and solutions and those who focus on individual work. The third factor is a bit more diffuse but refers to giving students time to summarize and reflect on work and to persevere in solving mathematics problems.

Table 8-1 **Mathematics teaching strategies**

Factors	To what extent do you use the following instructional strategies?
Sharing/grouping	• Students share solutions to problems and investigations • Students work with concrete materials or manipulatives • Teach through problem solving and investigations • Students work in collaborative groups
Individual	• Students work individually on problems • Explain, demonstrate, and provide examples • Provide time for practice
Summarize/Reflect (–)	• Students summarize what was learned • Allow time for student reflection • Encourage students to persevere
Note: (–) indicates factors that showed negative loadings of their items during the factor analysis.	

Charts 8-1 to 8-3 give the mean factor scores on each of these strategies by jurisdiction and language. For the sharing/grouping strategy, three population clusters may be identified. Six populations, from Saskatchewan French to Newfoundland and Labrador on the graph, are higher than the Canadian average and higher than most other populations. Seven others, from Saskatchewan English to Manitoba English, are close to the Canadian average and different from the top six. Finally, Quebec English and French are below the Canadian average and below all others (except Yukon, where the error is relatively large).

For individual strategies, Yukon stands out as higher than the Canadian average and higher than almost all others. Newfoundland and Labrador is also significantly higher than the Canadian average. Most of the other populations cluster around the Canadian average. Alberta French and British Columbia French stand out as significantly below the Canadian average and below all others, keeping in mind that relatively few teachers reported in each of these populations.

On the summarizing/reflecting strategy, there is a wide spread with seven populations, from Yukon to Manitoba English on the graph, above the Canadian average, and eight populations, all francophone, below.

CHART 8-1 **Mean factor scores for teacher use of sharing/grouping instructional strategies by jurisdiction and language**

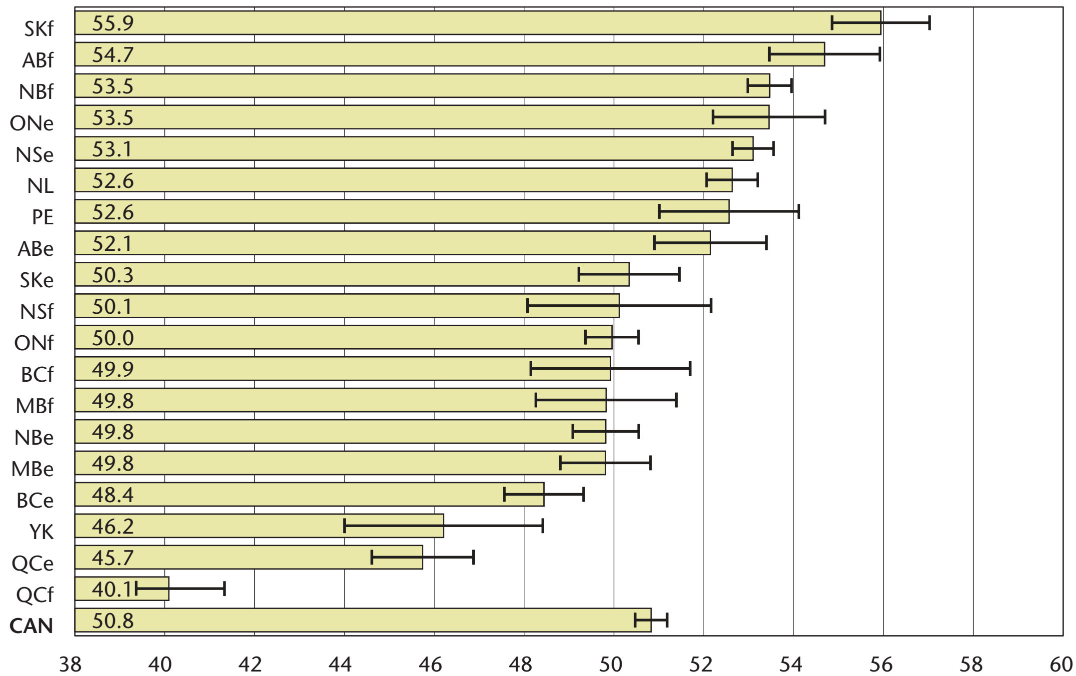


CHART 8-2 **Mean factor scores for teacher use of individual instructional strategies by jurisdiction and language**

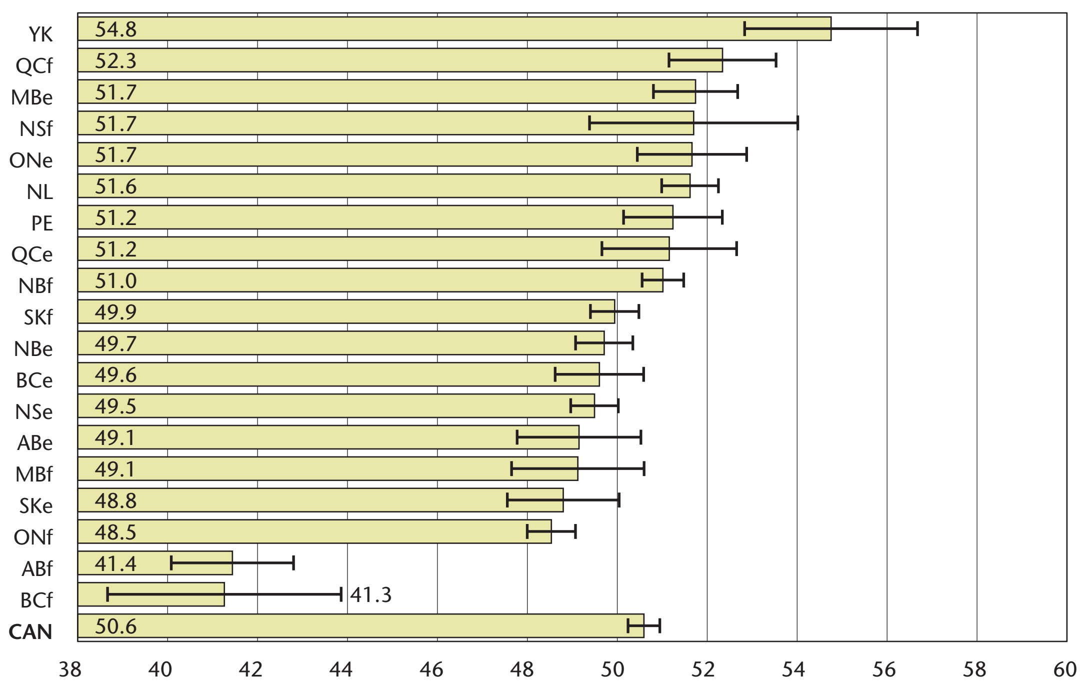


Chart 8-3 **Mean factor scores for teacher use of summarize/reflect instructional strategies by jurisdiction and language**

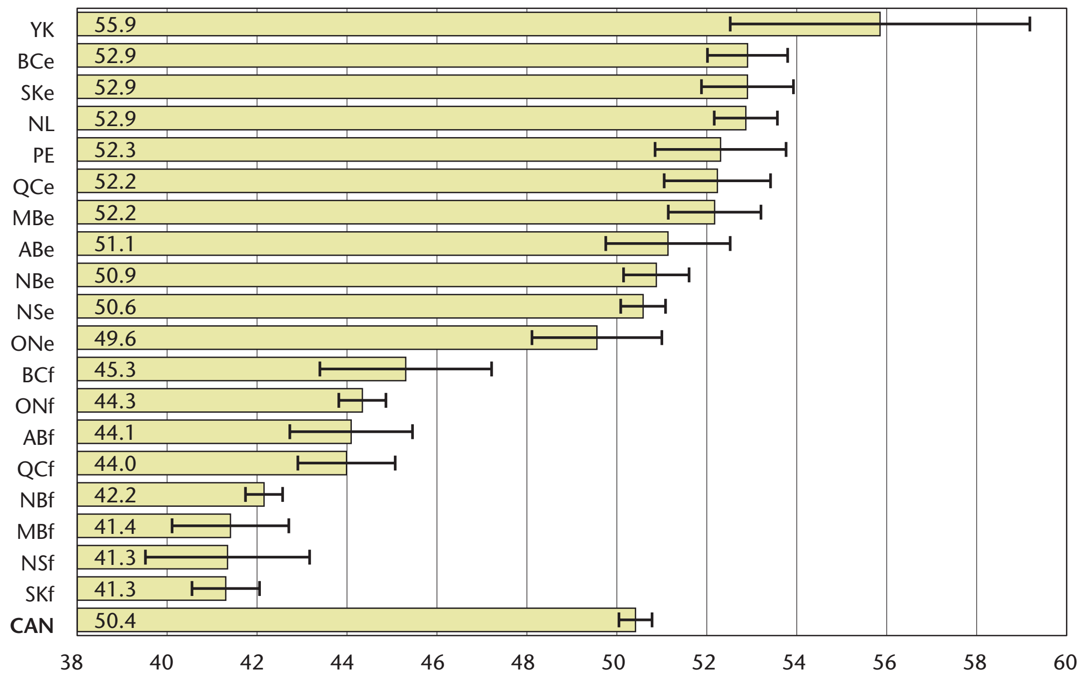
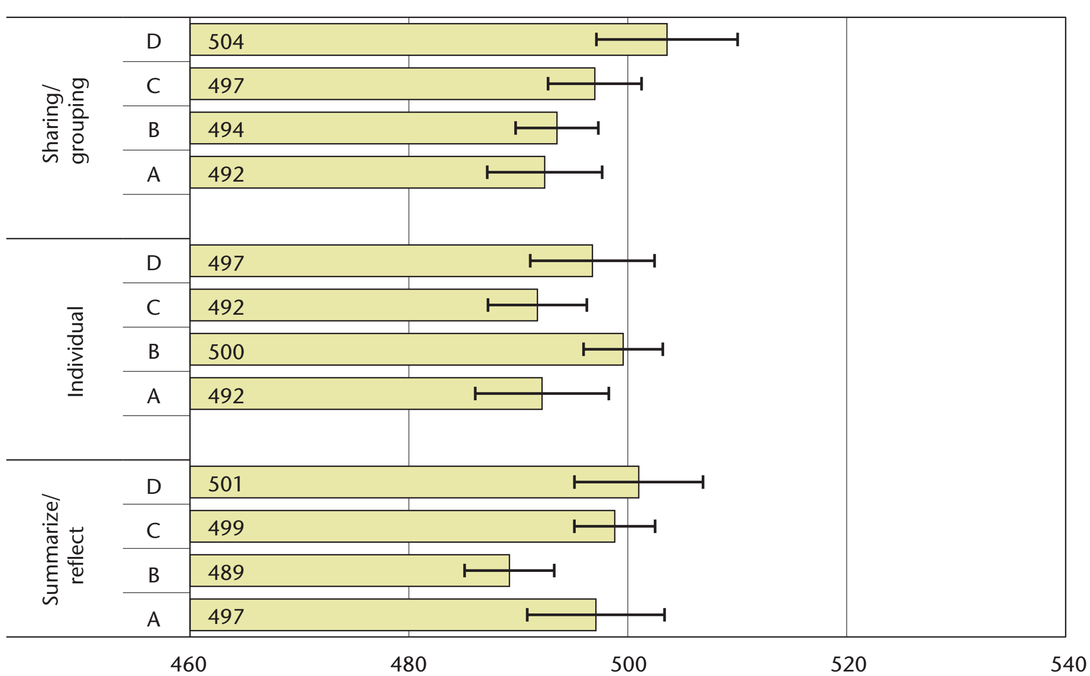


Chart 8-4 gives the mean mathematics scores for the four categories, each based on one standard deviation change in these teaching strategies. The pattern for sharing/group strategies is generally negative, though none of the differences is statistically significant. None of the other strategies shows any clear pattern.

Chart 8-4 **Mean mathematics scores by teaching strategies**



Mathematics learning strategies

A 9-item 3-point-value scale (slightly valuable, somewhat valuable, very valuable) was used to measure teacher views of the value of various mathematics learning strategies. This gave a three-factor solution as shown in Table 8-2. The first factor is interpreted as focusing on group work and the use of manipulatives. The second factor identifies the use of computers and calculators as distinct strategies. The third is labelled as a practice factor from the item with highest loading, with two other items, related to presenting problems and solutions in more than one way, being treated as variations on the practice theme.

TABLE 8-2 **Teacher views of the value of mathematics learning strategies**

Factors	In your view, how valuable are the following in helping students learn mathematics?
Groups/manipulatives	• Manipulatives • Class discussions • Working in groups • Presenting alternative methods of finding solutions
Computers/calculators	• Calculators • Computer software
Practice	• Practising • Presenting a concept in various ways • Presenting alternative methods of finding solutions

Charts 8-5 to 8-7 give the mean factor scores for these variables. Most teachers in the francophone populations are on the high end of the distribution for the groups/manipulatives factor. Quebec French is a notable exception to this pattern, standing out as lower than any others. For the value of computers/calculators, four francophone populations, along with Yukon and Saskatchewan English, are above the Canadian average. British Columbia French, Nova Scotia French, Newfoundland and Labrador, British Columbia English, and New Brunswick English are below. Finally, for the practice factor, six francophone populations, along with Prince Edward Island, New Brunswick English, and Newfoundland and Labrador, are below the Canadian average, and only one population, Quebec English, is above.

Chart 8-5 **Mean factor scores for teacher views of the value of groups/manipulatives learning strategy by jurisdiction and language**

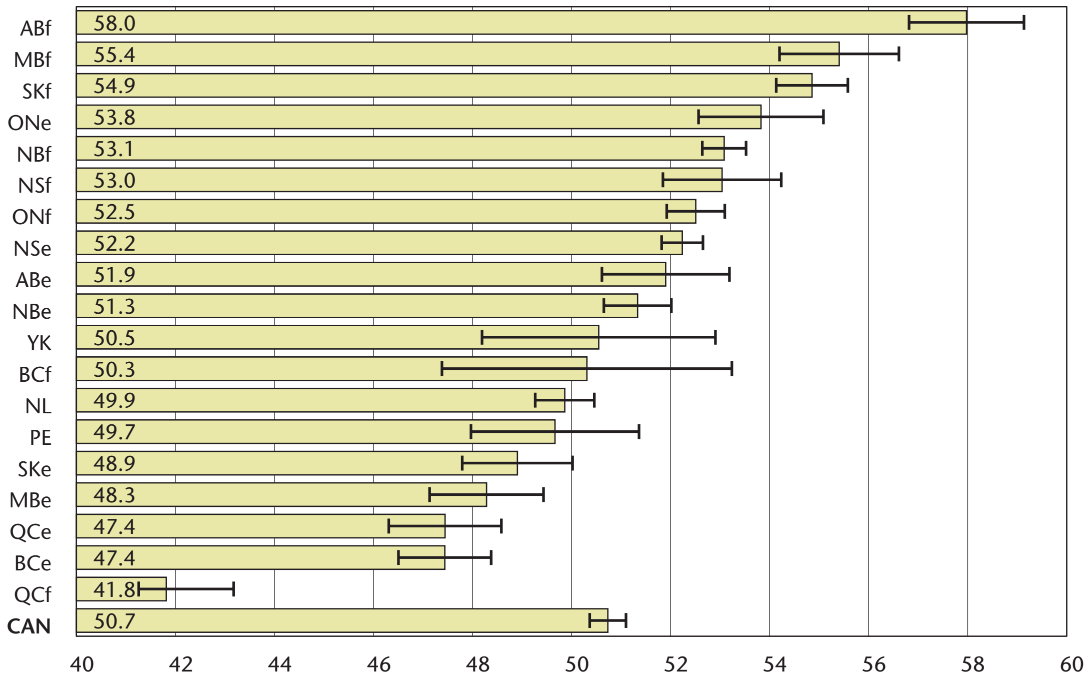


Chart 8-6 **Mean factor scores for teacher views of the value of computers/calculators learning strategy by jurisdiction and language**

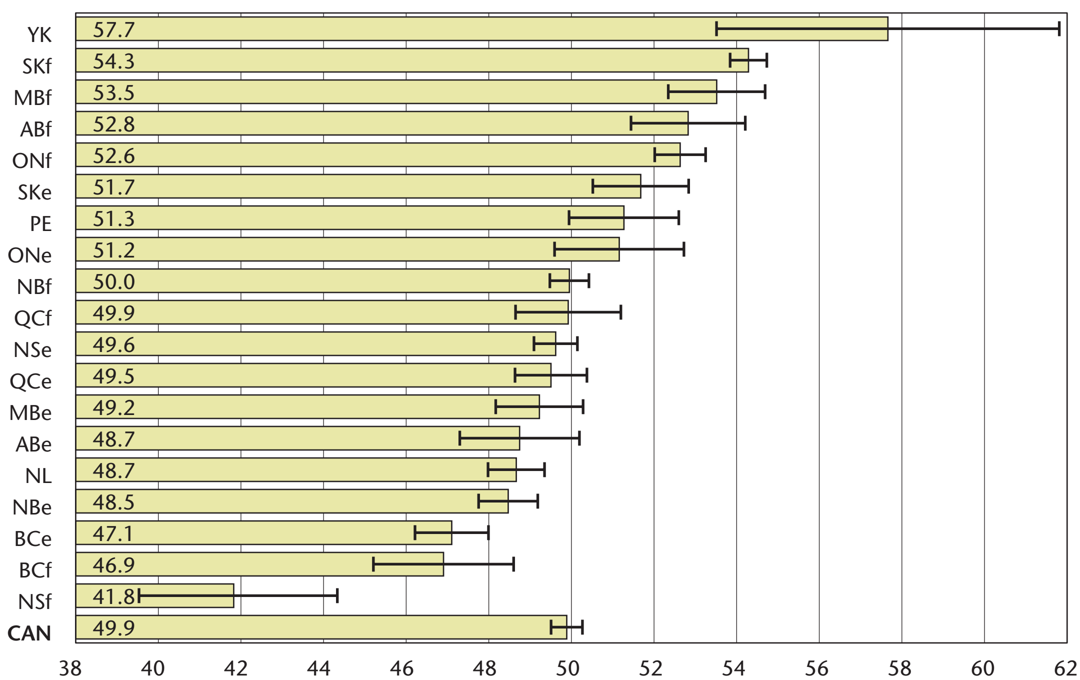


CHART 8-7 **Mean factor scores for teacher views of the value of practice learning strategy by jurisdiction and language**

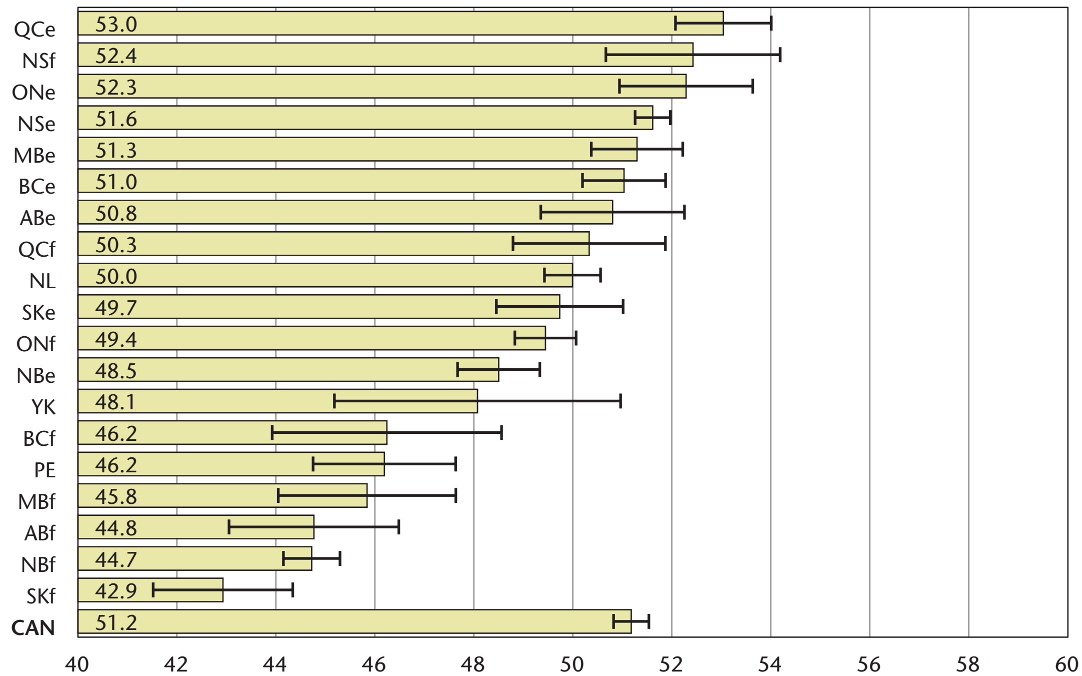
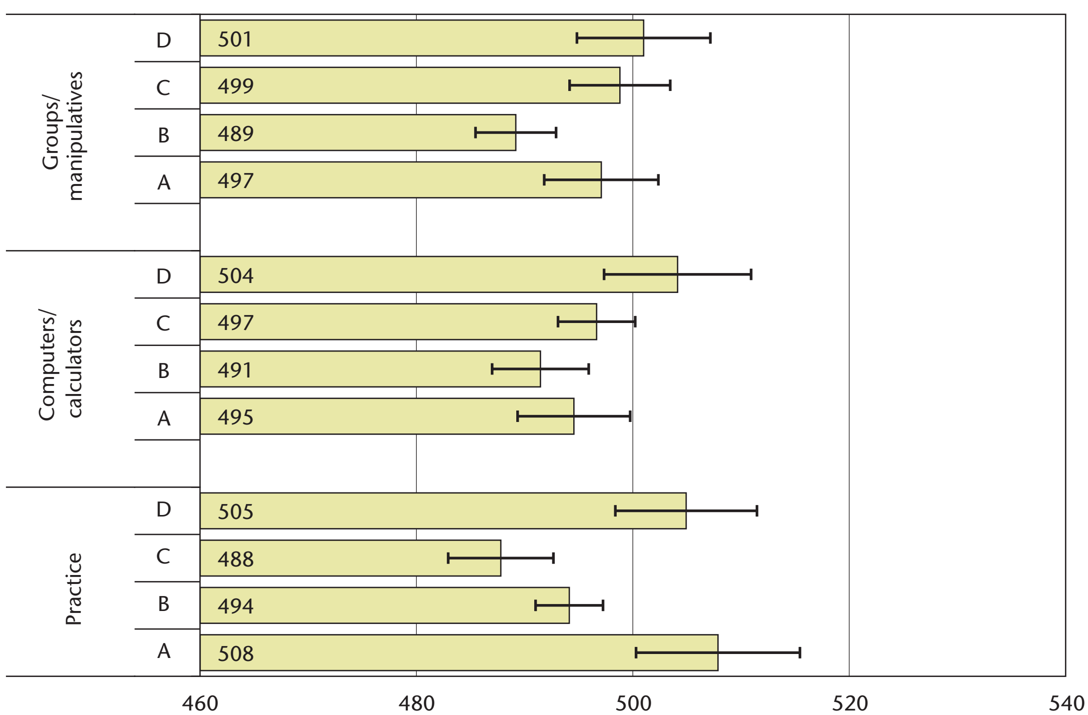


Chart 8-8 shows that these factors have little impact on mathematics scores. For the group effect, those in category B have a significantly lower mean score than those in categories D and C. However, there is no clear pattern beyond this. Practice shows a significant effect, with a non-linear pattern of higher scores for teachers at both extremes of the practice distribution.

CHART 8-8 **Mean mathematics scores for teacher views of the value of mathematics learning strategies**



A second set of seven learning strategy items was designed to examine how students represent or explain their mathematics reasoning. This was based on a 3 point frequency scale (rarely or never, sometimes, often). This yielded a two-factor solution shown in Table 8-3. Loadings on the first factor have to do with multiple representations and generalizations. The second factor is associated with explanations and justifications.

Table 8-3 **Representation and explanation as mathematics learning strategies**

Factors	How often do students do the following in your mathematics classes?
Multiple representations	• Make connections among multiple representations • Use multiple representations • Make generalizations and conjectures
Explanations	• Explain their solutions in writing • Explain their solutions orally • Justify their reasoning • Use correct mathematical language

Mean factor scores by population for these two factors are given in Charts 8-9 and 8-10. For multiple representations, Manitoba French, Nova Scotia English and French, and Ontario French are higher than the Canadian average; British Columbia French, along with Quebec English and French, is at the bottom. For the explanations factor, four francophone populations, along with Ontario English, are higher than the Canadian average, while British Columbia French, Nova Scotia English, Newfoundland and Labrador, Manitoba English, British Columbia English, and Saskatchewan English are below.

Chart 8-11 gives the mean mathematics scores for intervals of one standard deviation on these two variables. In both cases, the trend is toward higher mathematics scores for those who use both multiple representations and explanations more often. For multiple representations, those in Category A (more than one standard deviation above the mean) have significantly higher mathematics scores than those at the other three levels. For explanations, the differences are significant for levels D, C, and B, but not for A compared to B.

CHART 8-9 **Mean factor scores for multiple representations by jurisdiction and language**

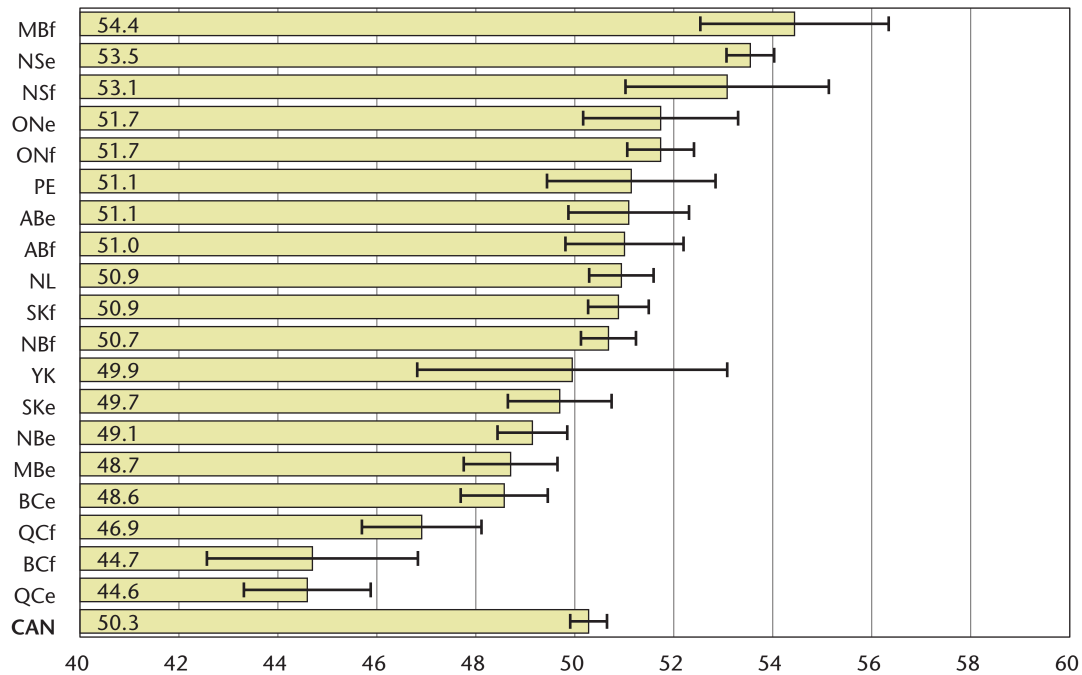


CHART 8-10 **Mean factor scores for explanations by jurisdiction and language**

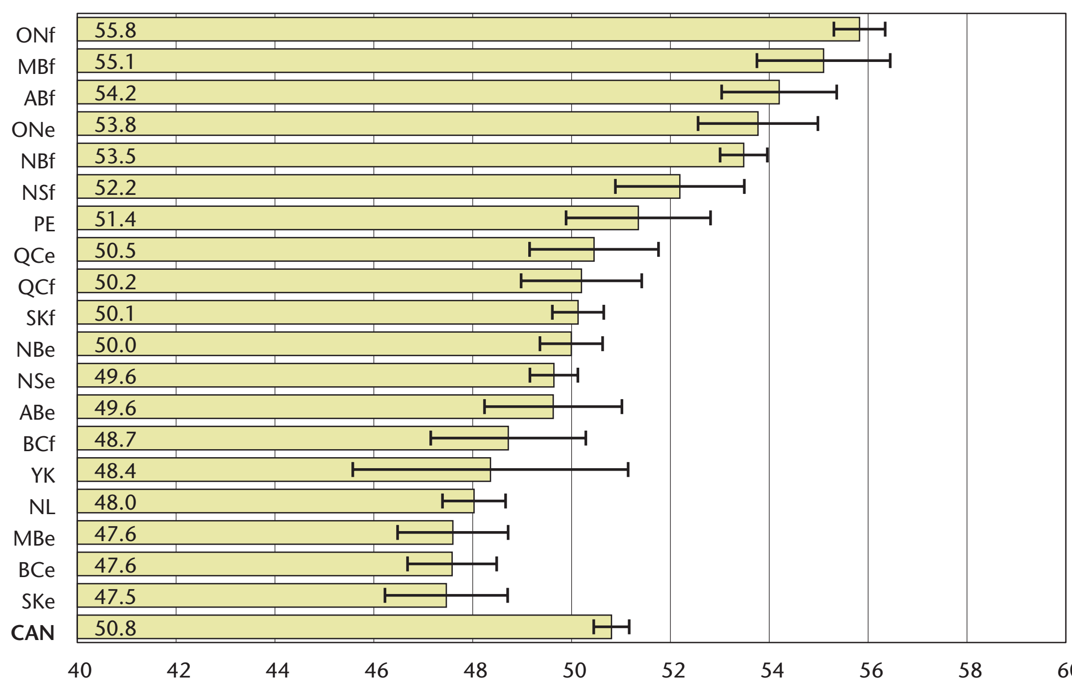
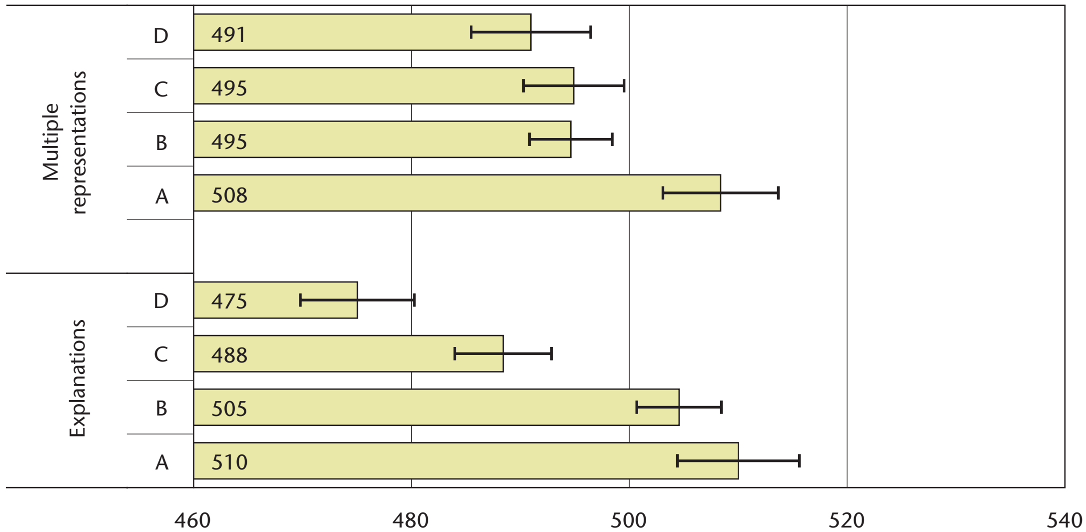


Chart 8-11 **Mean mathematics scores for multiple representations and explanations**



Learning resources

A 12-item 3-point frequency scale (rarely or never, sometimes, often) was used to measure the frequency of teacher use of various teaching and learning resources. This yielded four factors. The first three have fairly straightforward interpretations as indicated in Table 8-4. The fourth consisted of a mix of items that had no clear interpretation. This factor was thus not included as a derived variable.

Charts 8-12 to 8-14 show the mean factor scores for these variables. For technology, Alberta French and Ontario French are above the Canadian average and above all other populations. New Brunswick French, Alberta English, Ontario English, and New Brunswick English are also above the Canadian average. At the other extreme, British Columbia English is lower than the Canadian average and lower than all others. Nova Scotia English and Quebec English and French are also below the Canadian average. For print resources, 10 populations, from Quebec French to British Columbia French on the chart, are above the Canadian average. This group also included six of the eight French populations. Five populations, from British Columbia English to Alberta English on the chart, are below the Canadian average. For text resources, 10 populations, from Saskatchewan English to Ontario French, are above the Canadian average. Only two populations, Manitoba English and British Columbia English, are below.

Table 8-4 **Teacher use of mathematics learning resources**

Factors	How often do you use the following learning resources in your mathematics classes?
Technology resources	• Computer software • Web-based resources • Smart-boards • Spreadsheets • Measuring devices
Print resources	• Other print resources • Worksheets
Text resources	• Textbooks • Teacher's guides • Mathematics curriculum documents

Chart 8-12 **Mean factor scores for frequency of teacher use of technology learning resources by jurisdiction and language**

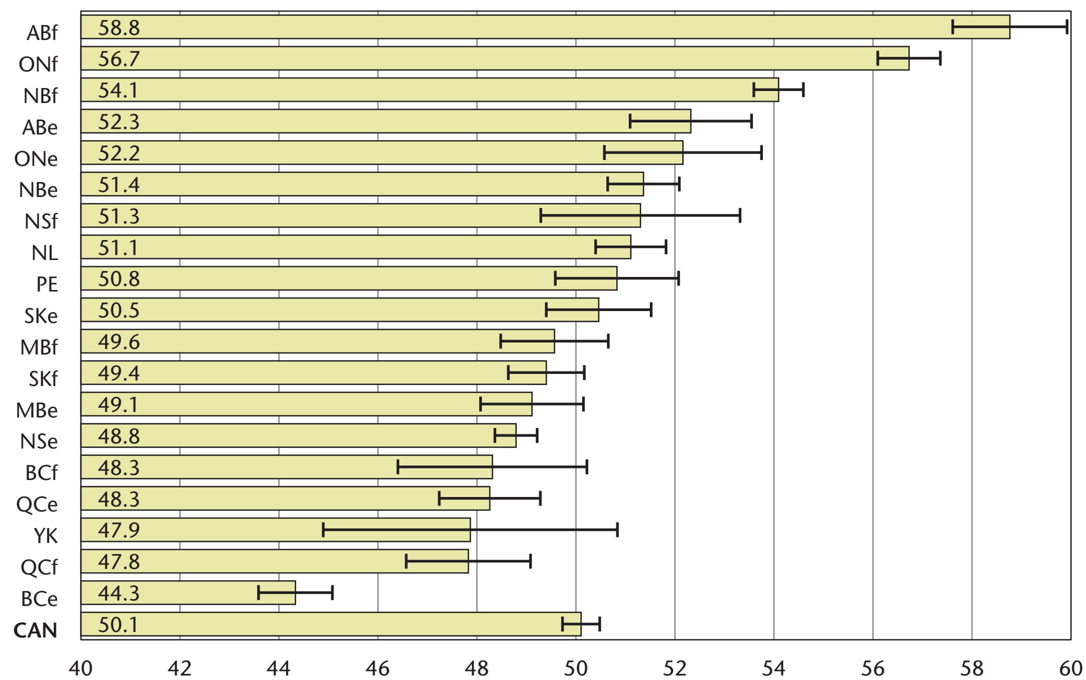


CHART 8-13 **Mean factor scores for frequency of teacher use of print learning resources by jurisdiction and language**

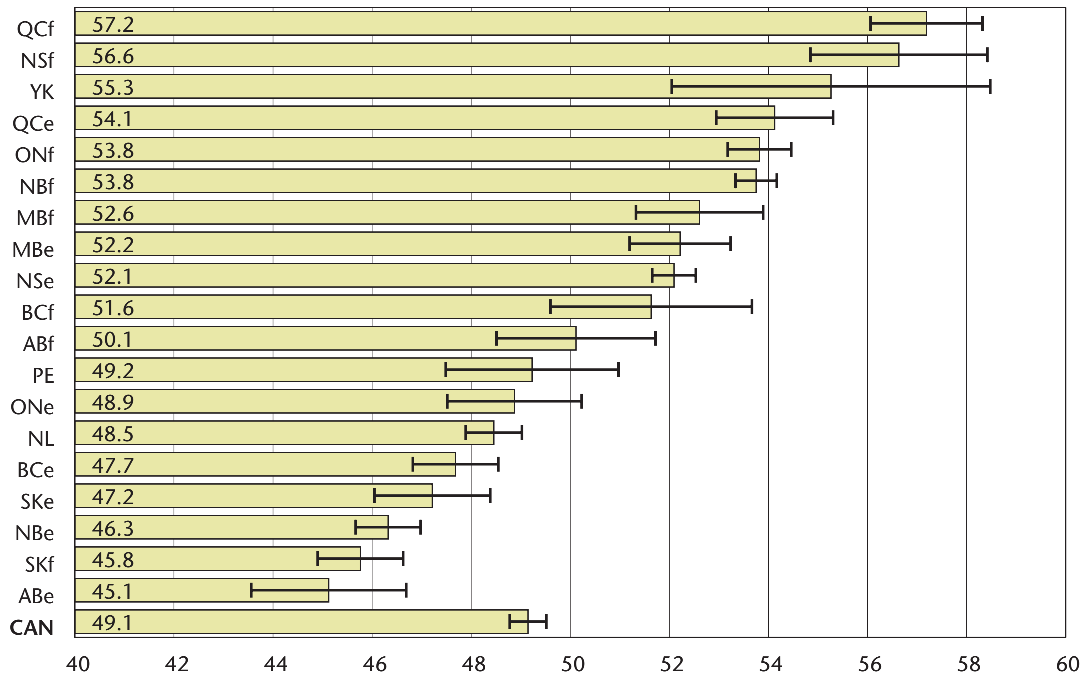


CHART 8-14 **Mean factor scores for frequency of teacher use of text learning resources by jurisdiction and language**

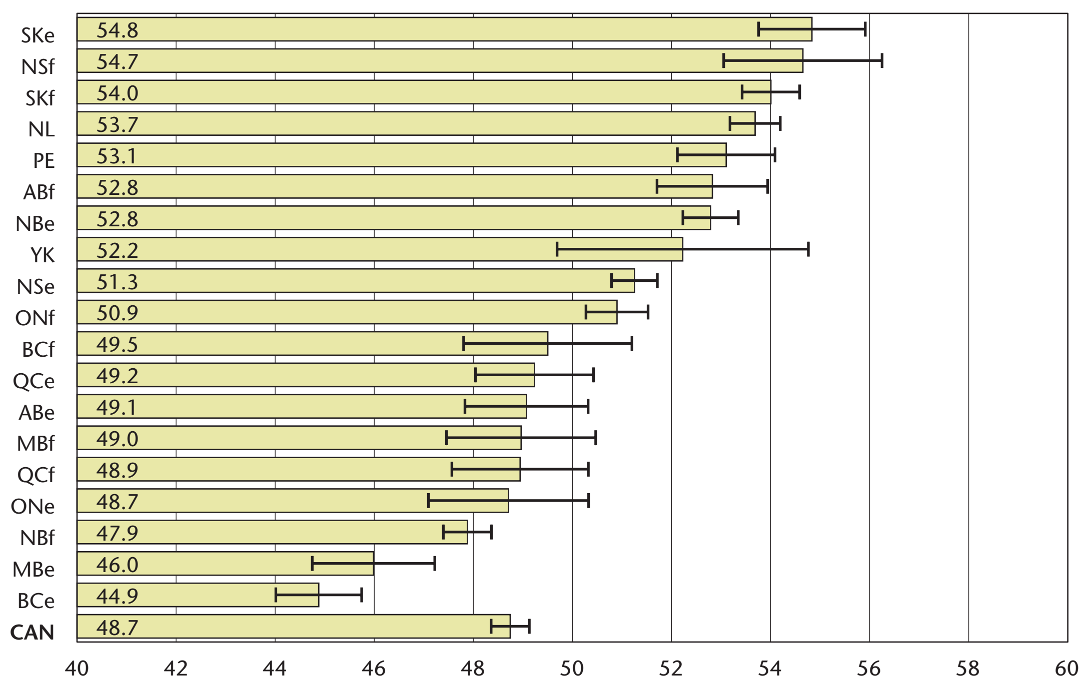
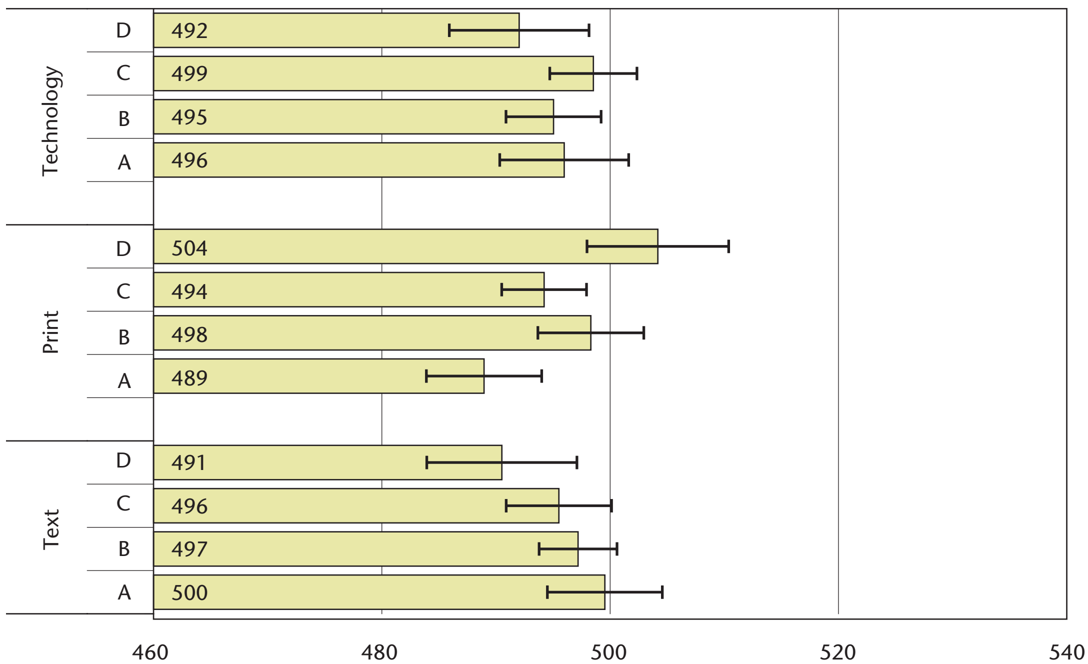


Chart 8-15 gives the mean mathematics scores for one standard deviation change in these variables. There is no clear pattern for use of technology resources. For print resources there is a significant difference between the lowest and highest category, with the latter showing a lower mathematics score. The pattern for text resources is in the direction of higher scores for more use of these resources; however, the differences between categories are not statistically significant.

Chart 8-15 **Mean mathematics scores for teacher use of mathematics learning resources**



Student reports of teaching strategies

Teaching strategies were further investigated by asking students a series of 17 questions on a 3-point frequency scale (rarely or never, sometimes, often). This gave the four-factor solution shown in Table 8-5. The clearest division is between the factors labelled direct instruction and indirect instruction. Explaining and justifying is a separate factor and finally the "use calculators" question stands out as one with a single high loading and is thus treated as a separate factor.

Table 8-5 **Student reports of teaching strategies**

Factors	How often do you do the following in your mathematics classroom?
Direct instruction	• Watch the teacher do examples • Listen to the teacher give explanations • Copy notes given by the teacher • Practise new skills • Teacher-guided investigations • Review skills learned • Solve problems • Work individually on investigations of problems
Indirect instruction	• Use manipulatives (e.g., base-ten blocks, colour tiles, geometric solids) • Use computer software • Work in groups on investigations or problems • Share solutions with other students and with the class • Have opportunities to reflect on what was learned
Explanation/justify (–)	• Justify my reasoning • Explain my answers • Use my own strategies to solve problems
Use calculators	• Use calculators
Note: (–) indicates factors that showed negative loadings of their items during the factor analysis.	

Charts 8-16 to 8-19 show the mean scores on each of these factors by jurisdiction and language. For direct instruction, Newfoundland and Labrador, British Columbia English, Quebec English, and Nova Scotia French are above the Canadian average. Five of the other francophone populations, along with Saskatchewan English, Nova Scotia English, and New Brunswick English are below the Canadian average. For indirect instruction, most populations are above the Canadian average, with Alberta French and Saskatchewan French also above all others. Nova Scotia French and Quebec English and French are below the Canadian average, with Quebec French standing out as below all others. The explanation/justification factor shows the opposite effect, with most populations below the Canadian average, with only Quebec English and Ontario French above. Finally, there is wider variation in the use of calculators than on the other factors. Most populations are below the Canadian average on this factor, while five populations, including Quebec English and French and three other francophone populations, are above.

Chart 8-16 **Mean factor scores for frequency of student use of direct instruction by jurisdiction and language**

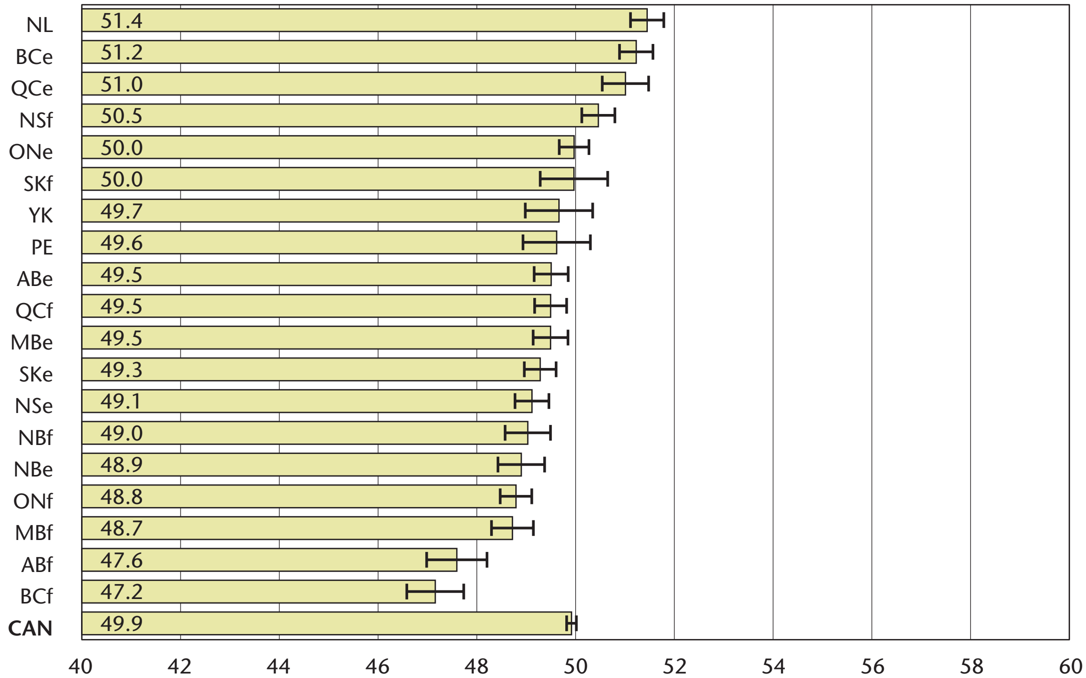


Chart 8-17 **Mean factor scores for frequency of student use of indirect instruction by jurisdiction and language**

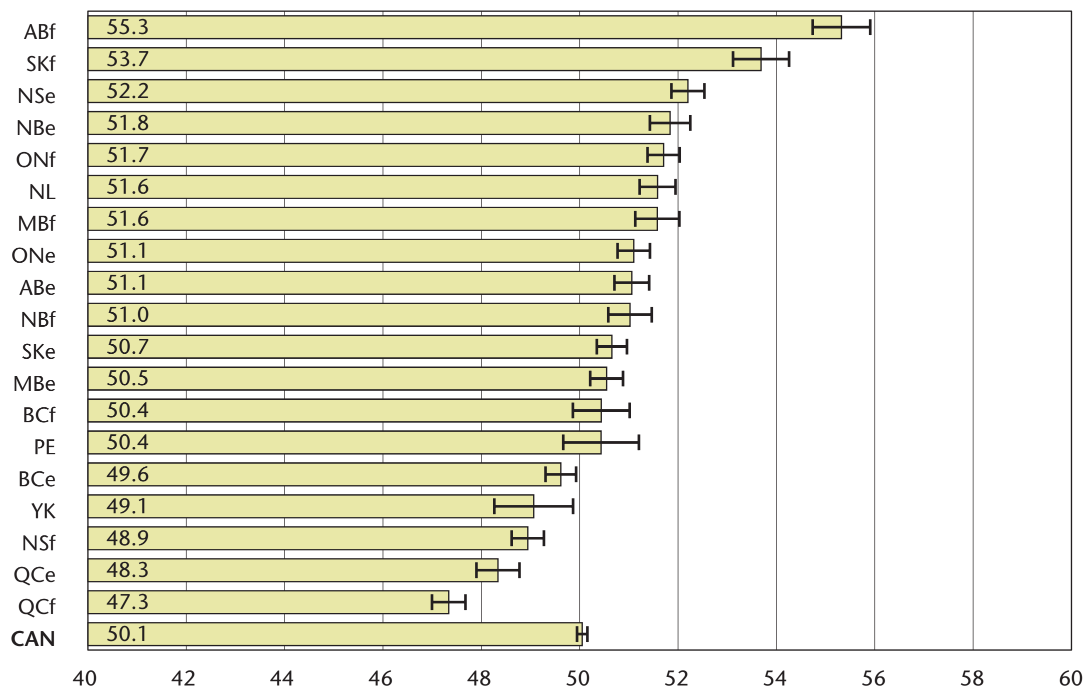


CHART 8-18 **Mean factor scores for frequency of student use of explain/justify by jurisdiction and language**

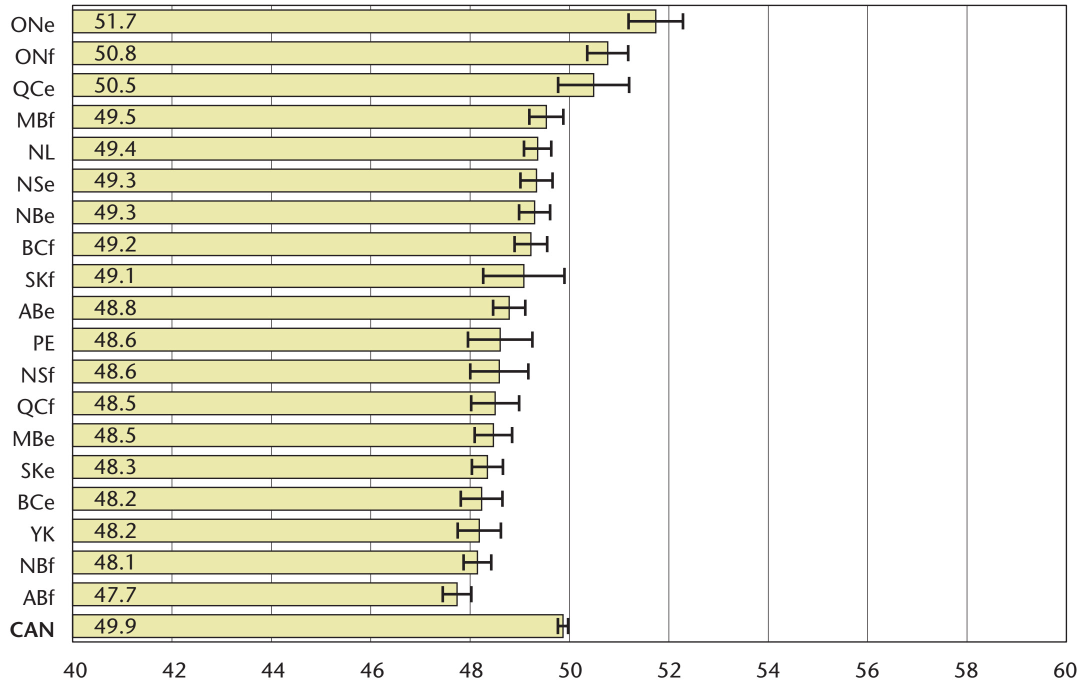


CHART 8-19 **Mean factor scores for frequency of student use of calculators by jurisdiction and language**

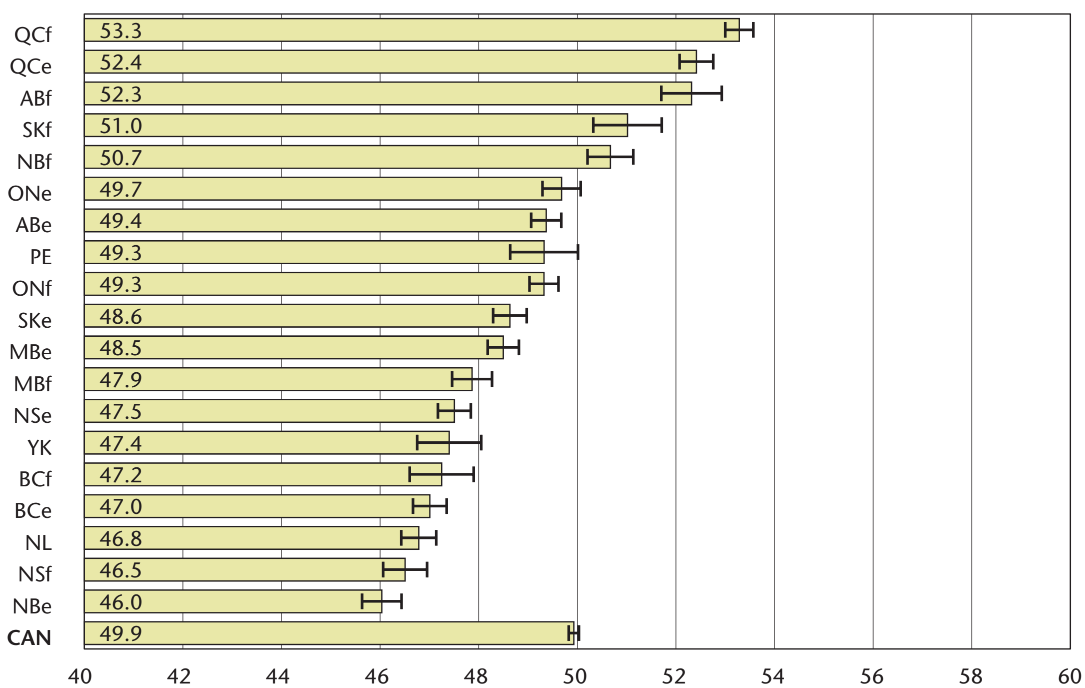
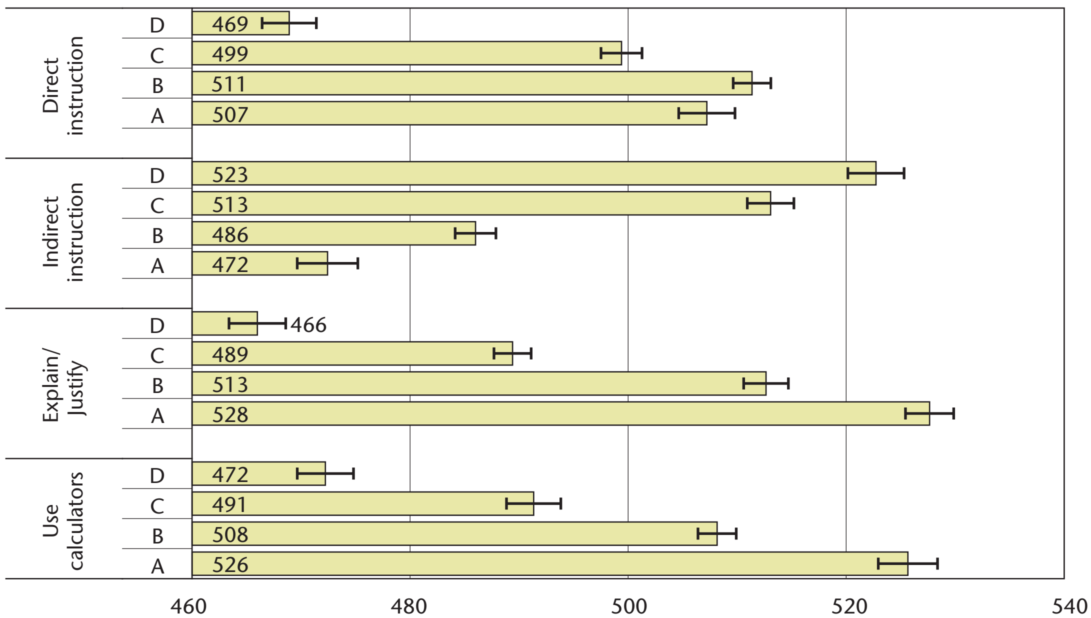


Chart 8-20 shows the mean mathematics scores for each of these factors. Most of these effects are highly significant. Greater use of direct instruction is associated with higher scores, except for the highest category, which is not significantly different from the second highest. Greater use of indirect instruction is strongly associated with lower scores. Greater use of explanation/justification is associated with higher scores, as is the use of calculators.

Chart 8-20 **Mean mathematics scores for student-reported instructional strategies**



Mathematics assignments

Students were asked about the types of mathematics assignments they are asked to complete, using a three-point frequency scale (rarely or never, sometimes, often). This gave two factors as shown in Table 8-6. There seems to be a clear distinction between classes for which assignments are project or group-based and those that are textbook-based. Use of worksheets has a negative loading on the latter factor, suggesting that those who make higher use of textbooks make lower use of worksheets.

Charts 8-21 and 8-22 give the mean factor scores on these two factors by population. On the projects factor, six populations, from Alberta French to Ontario English on the chart, are significantly above the Canadian average, and eight, from New Brunswick English to Nova Scotia French, are significantly below. For textbook assignments, almost all populations are either significantly above (from Newfoundland and Labrador to British Columbia English) or significantly below (from Saskatchewan French to New Brunswick French) the Canadian average. Only two, Ontario English and Nova Scotia English, are at the Canadian average.

TABLE 8-6 **Mathematics assignments**

Factors	How often do you have the following kinds of assignments in your mathematics classes?
Project assignments	• Group projects requiring work outside of class • Individual projects requiring work outside of class • Group work in the classroom
Textbook assignments	• Questions from textbooks • Worksheets (–)
Note: (–) indicates items that showed negative loadings during the factor analysis.	

CHART 8-21 **Mean factor scores for project assignments by jurisdiction and language**

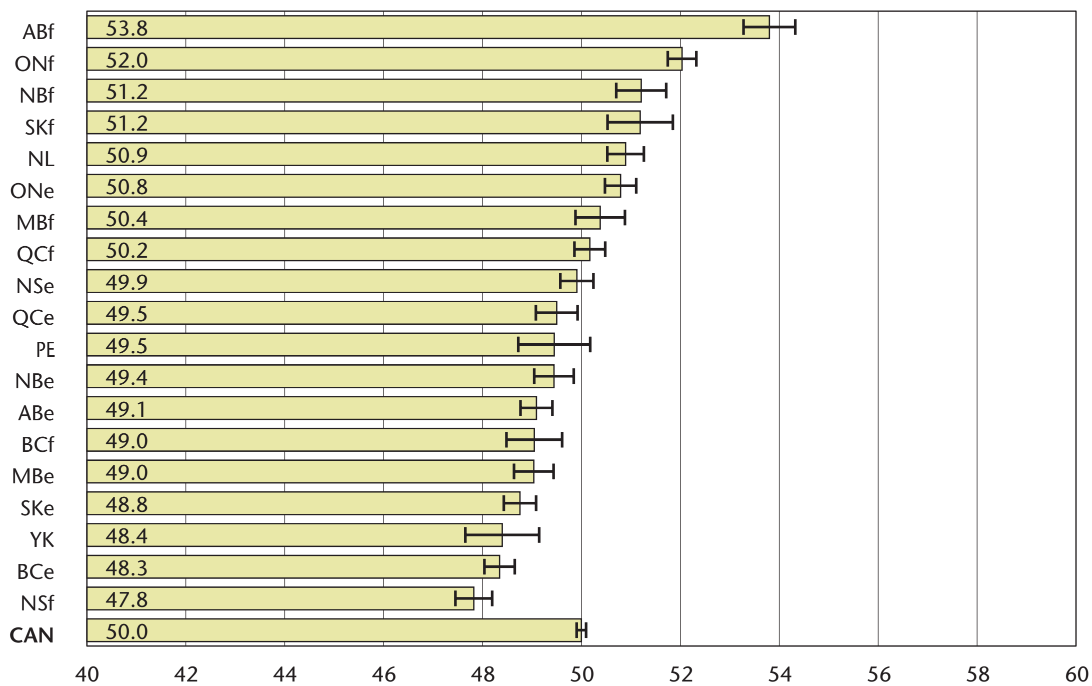


Chart 8-22 **Mean factor scores for textbook assignments by jurisdiction and language**

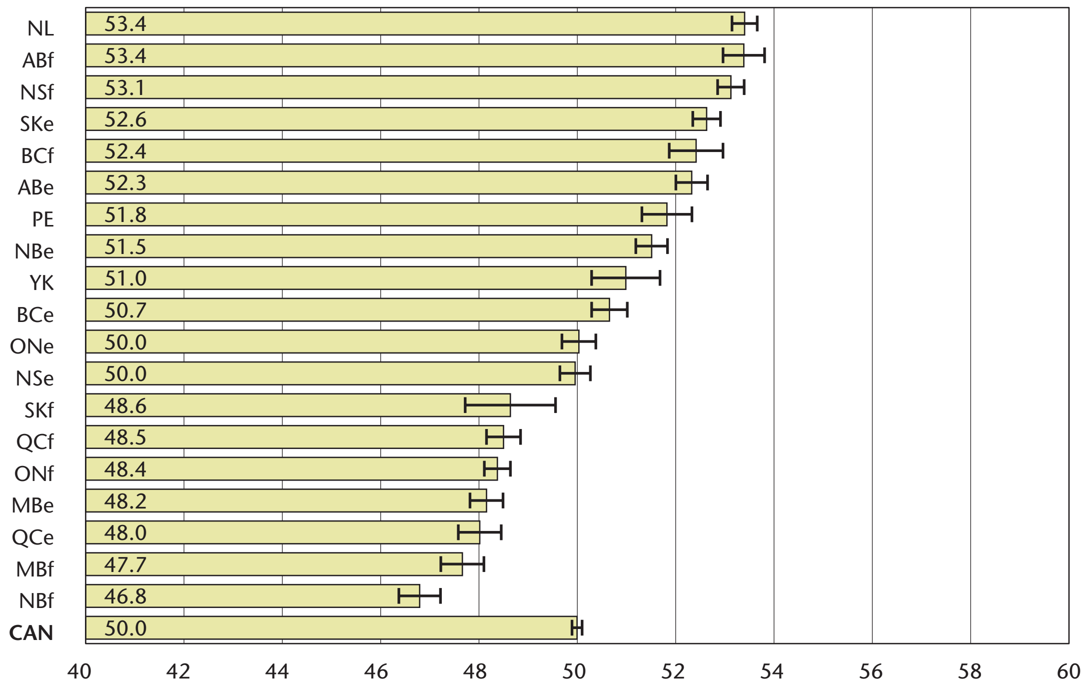
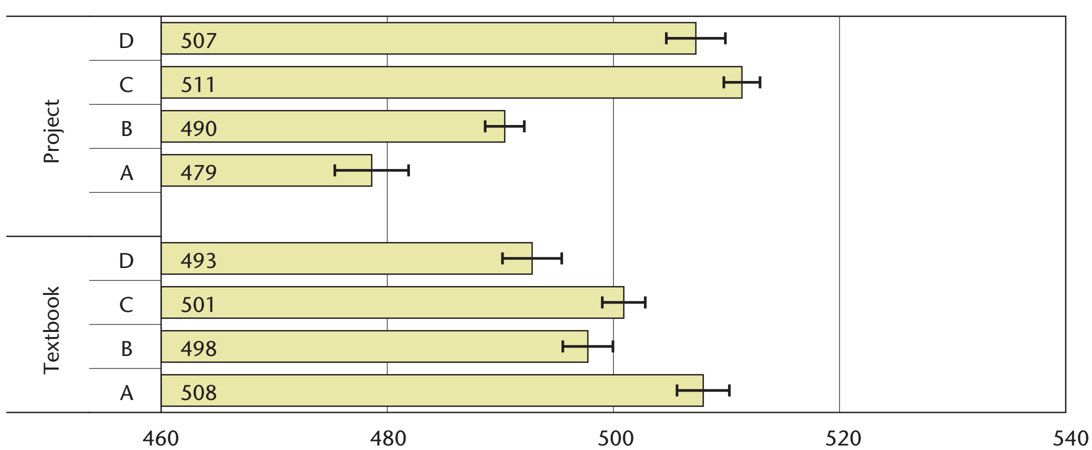


Chart 8-23 gives the mean mathematics scores for standard deviation unit changes in these two variables. Although the trend is not strictly linear, the general pattern is toward lower mathematics performance with greater use of project assignments, and higher mathematics performance with greater use of textbook assignments. It is noted that, since worksheet use has negative loadings on this factor, in the model, the results can be interpreted to mean that textbook use has a positive effect, but worksheet use a negative one on mathematics achievement.

Chart 8-23 **Mean mathematics scores for mathematics assignments**



Multiple regression effects

The eight variables representing teacher-reported strategies were analyzed separately and as a group using a two-level model as before. Resource uses were also analyzed in the same way but separately from teaching strategies.

Chart 8-24 shows relatively few of these effects to be statistically significant in either model. The only exception is the sharing/group strategy, which changes from non-significant to significantly negative in the multiple regression model. Using explanations has a significant positive effect on mathematics scores in both models.

CHART 8-24 **Regression coefficients for teacher-reported mathematics instructional strategies and resources**

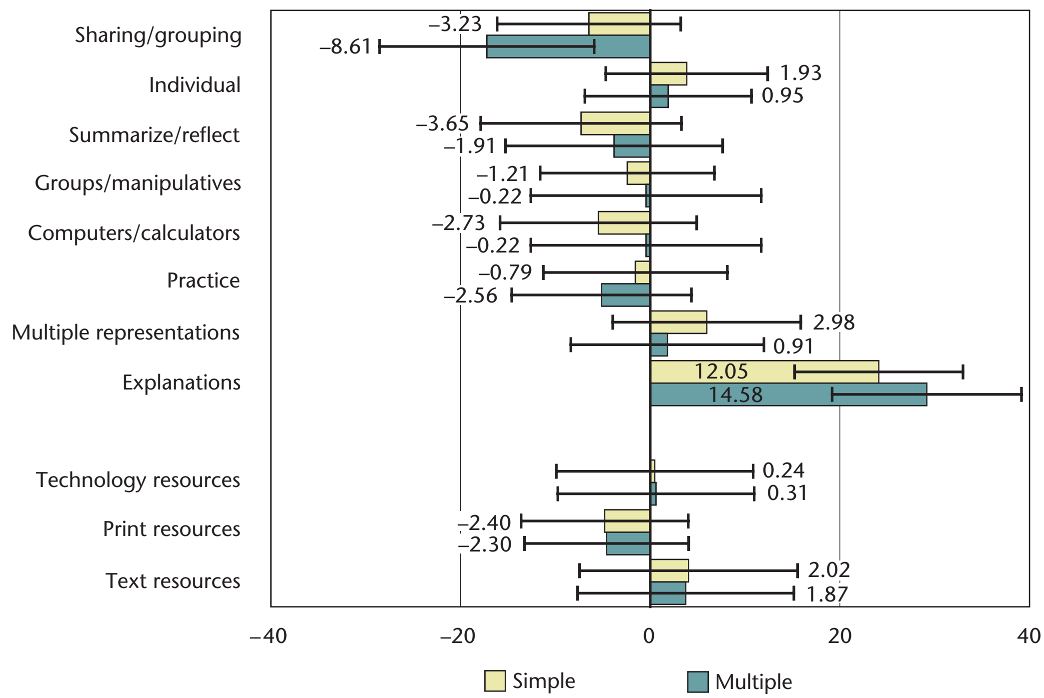
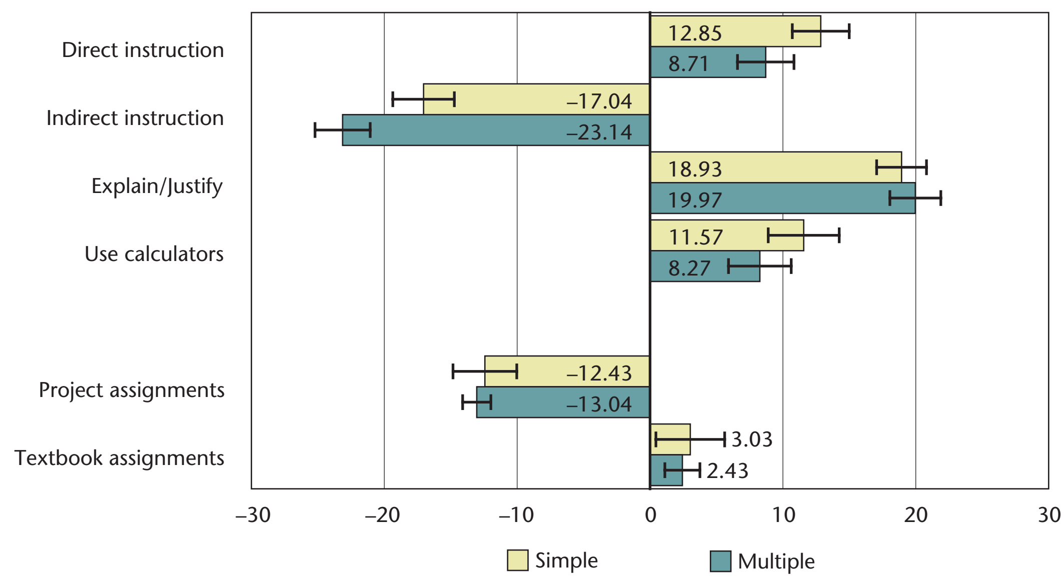


Chart 8-25 shows the effects of the student-reported instructional strategies and assignments. In contrast to the teacher effects, most of these are highly statistically significant, with only small changes from the simple to the multiple regression model. The results for the explain/justify factor are similar to those for the same factor in the teacher chart.

The main change is that direct instruction becomes significantly more positive and indirect instruction significantly more negative when other variables are controlled. This indicates that in a situation in which both strategies are used, controlling for one of these increases the effect, whether positive or negative, of the other. Thus, for example, some use of direct instruction can be said to offset the negative effects of indirect instruction.

Chart 8-25 **Regression coefficients for student-reported mathematics instructional strategies**



9 ASSESSMENT

Questions on assessment were included on each of the student, teacher, and school questionnaires. These had two main focal points. The first was to look at assessment practices used by schools and teachers and their impact on performance. The second was concerned with awareness, use and impact of large-scale assessment, including provincial/territorial assessments, SAIP, PCAP, and PISA. More specifically, questions or question sets were designed to examine classroom assessment methods, awareness and use of rubrics, types of items used in assessment, the use of non-academic criteria in grading, the availability and use of external assessments, and the purposes for which assessments are used.

Methods of classroom assessment

Students were asked how often they are assessed using each of eight different methods, using a three-point time scale (rarely or never, sometimes, often). Factor analysis of these items yielded three factors as shown in Table 9-1. The first encompasses a range of what may be termed "unconventional" assessments or those of more recent origin. The second groups together tests and homework, two relatively conventional methods of assessment. The third distinguishes exams from all other methods. If we interpret the term "exams" to encompass formalized end-of-term or end-of-year assessments, whether internal or external to the teacher or school, it can be seen that this can stand apart from the remaining two conventional methods of Factor 2.

Table 9-1 **Assessment methods**

Factors	In your mathematics classes, how often are you assessed using the following methods?
Unconventional assessment	• Self-assessment • Peer assessment • Journals • Portfolios • Group Work
Conventional assessment	• Test/quizzes • Homework
Exams	• Exams

Mean factor scores for these factors by jurisdiction and language are given in Charts 9-1 to 9-3. For unconventional assessment, two populations, Newfoundland and Labrador and Ontario English, are above the Canadian average and above all others. Two other populations, Manitoba English and Nova Scotia French, are above the Canadian average. Seven populations, Ontario French, Yukon, British Columbia English and French, Quebec English, Quebec French, and New Brunswick French, are below the Canadian average.

The conventional assessment variable shows a distinct cluster of five populations, all English, above the Canadian average and above all others. Most other populations are below the Canadian average, with a second cluster, all French, below all others.

For the exams variable, Quebec French stands out as using this method more than all others. Most other populations are above the Canadian average. Five populations, from Manitoba English to Ontario English on the chart, are below the Canadian average and below all others. With one exception, these are significantly different from each other, suggesting a wide spread on this variable among those with low use of exams for assessment.

Chart 9-1 **Mean factor scores for assessment with unconventional methods by jurisdiction and language**

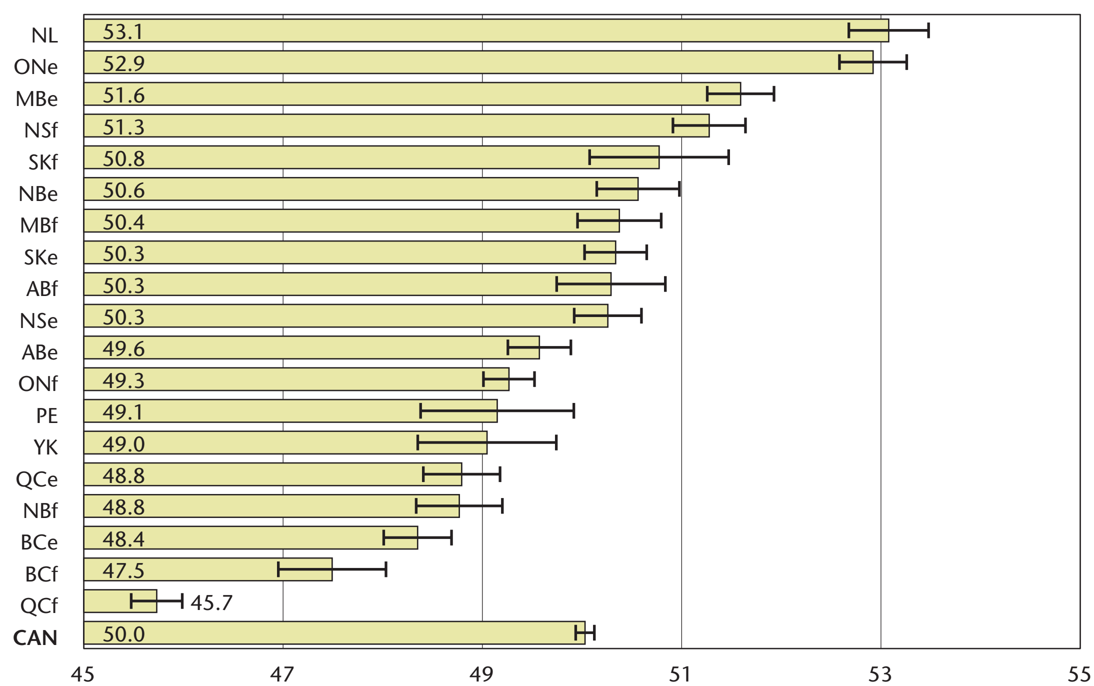


Chart 9-2 **Mean factor scores for assessment with conventional methods by jurisdiction and language**

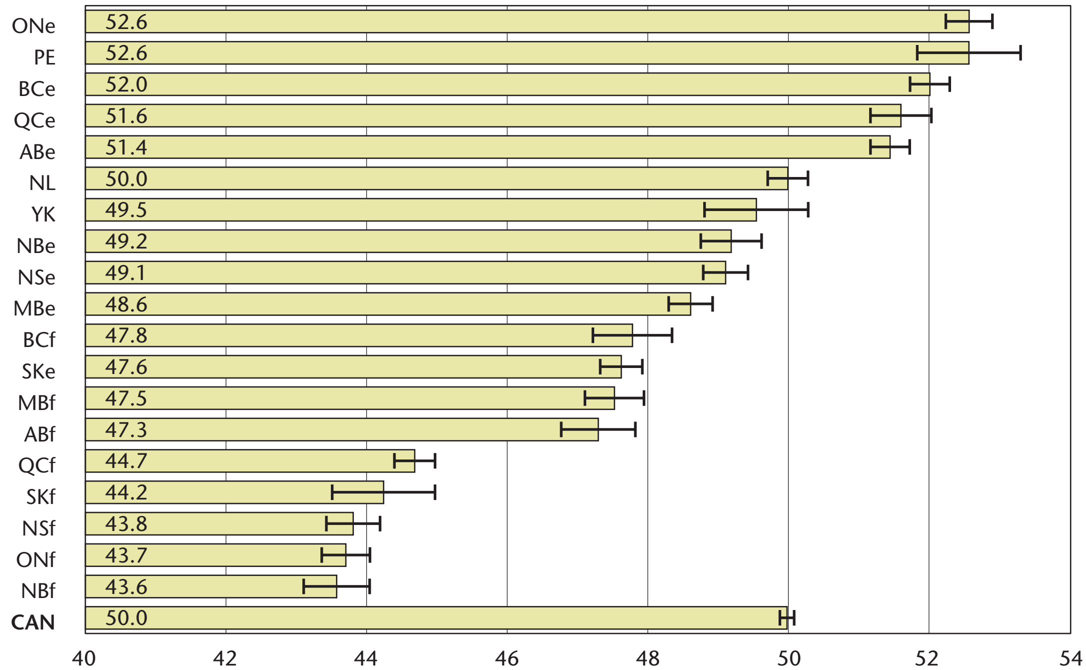
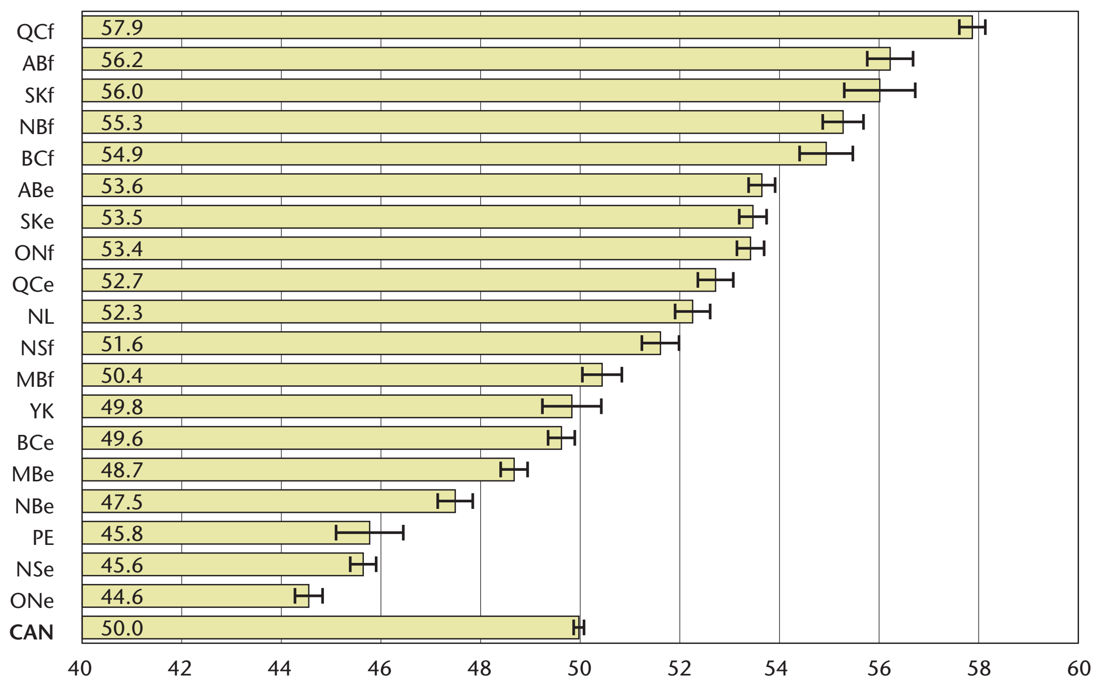
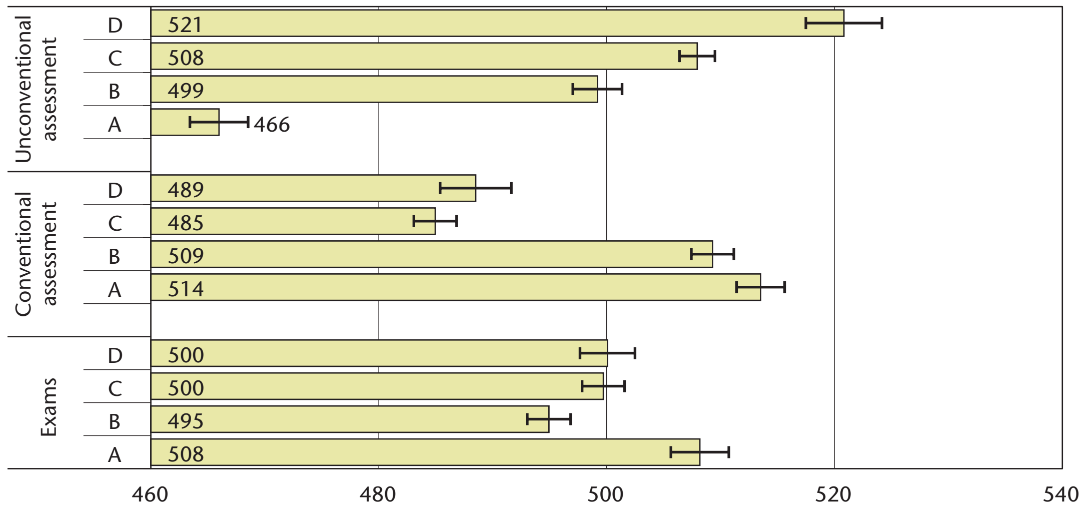


Chart 9-3 **Mean factor scores for assessment with exams by jurisdiction and language**



Mathematics mean scores for these assessment methods are given in Chart 9-4. Use of unconventional assessment methods is strongly negatively associated with mathematics achievement. Use of conventional assessment methods shows a positive effect, with a clear division between the two lowest categories (D and C) and the two highest (B and A). Exam use shows a less clear pattern, with the highest use being associated with the highest mean mathematics scores, the next highest use with the lowest scores, and about average achievement for those at the low end of the scale on the exams variable.

Chart 9-4 **Mean mathematics scores by assessment methods**



Rubrics

Rubrics are statements designed to capture the desired outcome, and the level of performance expected on that outcome, for a particular learning task. Rubrics are used mainly for scoring when the scoring criteria are qualitative. However, they may also be used to inform students of expectations and to structure learning tasks. Rubrics are now widely used in teaching as a means of clarifying outcomes and expectations. An example of a rubric is the table below, which describes what students can be expected to be able to do at each of the four proficiency levels of the PCAP mathematics scale.

Level 1 – Scores of 357 and less	Example
Students at this level were able to solve problems at a low cognitive level that were determined to be fairly easy questions. Typically, at this level, students were able to retrieve information from a graph or solve previously learned routine problems. At this level, students could solve problems that required mostly recall and recognition.	The person who delivers Martine’s meals to her customers charges her a fee for the deliveries as shown in the table below. Complete the table to show the total of the delivery charges for the week.

Monday	$32.75
Tuesday	$27.40
Wednesday	$41.95
Thursday	$38.05
Friday	$65.25
Saturday	$49.50
Sunday	$46.40
Total	

Level 2 – Scores between 358 and 513	Example
Students at this level were required to recall facts, definitions, or terms and carry out previously learned procedures such as performing one or more operations, employing formulae, evaluating a variable expression, retrieving information from a table or a graph and applying it to solve a problem. Typically, students at this level were able to identify a simple number of geometric patterns. Students were able to solve problems that were clearly defined as to what was required, with no extraneous information or hidden assumptions. At this level, students could solve problems that were mostly of low and moderate cognitive demand.	Mr. Robert rides his bike to school every day. He also uses his bike as a tool to teach his students a few concepts about circles.

What is the diameter of the front wheel of Mr. Robert’s bike?

A. 45 cm

B. 80 cm

C. 85 cm

D. 90 cm

Level 3 – Scores between 514 and 668	**Example**
Students at this level were able to apply what they know to new situations, identify hidden assumptions, and distinguish between relevant and irrelevant information needed to solve a problem. They had to select appropriate procedures or strategies to solve a problem and sometimes had to apply skills from different domains to solve problems. Students at this level were able to represent a problem in different ways and use informal reasoning to solve problems. At this level students could solve problems that were mostly of moderate to high cognitive demand.	 A talent show will start with a 10-minute introduction, and each skit is allowed 5 minutes. The talent show is scheduled to start at 7 p.m. and end at 9 p.m. The total length of time of the talent show can be represented by the equation $T = 10 + 5s$ where T represents the total time of the show in minutes, and s represents the number of skits. Using the equation, determine how many skits will be in the talent show. Show your work.
Level 4 – Scores at 669 and above	**Example**
Students at this level were able to solve problems that require complex reasoning at the analysis and synthesis levels. Solutions clearly show a mastery of the appropriate conceptual and procedural knowledge necessary to solve complex problems. Students were able to generalize a pattern and write the rule algebraically. They were also able to explain or justify their solutions and strategies clearly. At this level, students could solve problems that were generally of high cognitive demand and determined to be difficult questions.	Sarah plays a game. After two weeks, Sarah has 105 points. After the third week, she has 135 points. Which of the following could be used to calculate the percentage increase in Sarah's point total? A. $\frac{135-105}{135} \times 100$ B. $\frac{135-105}{105} \times 100$ C. $\frac{135}{105} \times 100$ D. $\frac{105}{135} \times 100$

The rubric used for determining communication proficiency levels and examples of student work is shown below.

Code	Student's exemplars
Code 3 Code description: There is a clear description of the student's reasoning, with a logical, organized, and precise use of mathematical procedures, notation, and proper labelling. Rationale: For this item the response had to be clearly labelled, with logical work that justified the answer. In the following example the student has shown an explicit conversion, there are no skipped steps, and the units are included in the answer.	**Show your work.** $9 - 7 = 2$ 2 hours = 120 minutes $5\overline{)110}$ = 22 $120 = 10 + 5s$ $120 - 10 = 10 + 5s - 10$ $110 = 5s$ $110/5 = 5s/5$ $22 = s$ **Answer:** 22 skits
Code 2 Code description: There is an adequate description of the student's reasoning to arrive at the answer given. Rationale: The work illustrated the steps taken but had minor elements missing. In the following example, the student did not show where he or she obtained the value of 120.	**Show your work.** $T = 120$ $T = 10 + 5s$ $120 = 10 + 5s$ $120 - 10 = 5s$ $110 = 5s$ $\frac{110}{5} = s$ $22 = s$ **Answer:** 22 skit
Code 1 Code description: There is a description of the student's reasoning, but the coder must make major assumptions or fill in major gaps. Rationale: In this example, there is no explanation for the 120, there are no units in the answer, and there is incorrect notation (incorrect use of the equal sign), but the coder can still follow the student's reasoning.	**Show your work.** $T = 10 + 5s$ $120 = \quad s = \frac{110}{5} = 22 + 10$ **Answer:** 22
Code 0 Code description: An answer, but with little or no communication of the process used.	**Show your work.** 10 5 7:10 1:20 mins hour $120 \div 5 =$ **Answer:** 24

Overall, about two-thirds of students indicated that they knew what a rubric is, and about half of those reported that rubrics are sometimes used in their mathematics classes. Detailed responses by jurisdiction and language are given in Chart 9-5. It is clear from this that students in most francophone jurisdictions are less familiar with rubrics and use them less than those in anglophone jurisdictions. The exception is Ontario French. The gap between knowing and use is also larger in some of the francophone jurisdictions. This is especially true in Ontario where knowledge is relatively high in both populations, but use at the start of assignments is much lower among francophones than among anglophones.

CHART 9-5 **Percentage of students who know what a rubric is and who sometimes use a rubric in mathematics classes by jurisdiction and language**

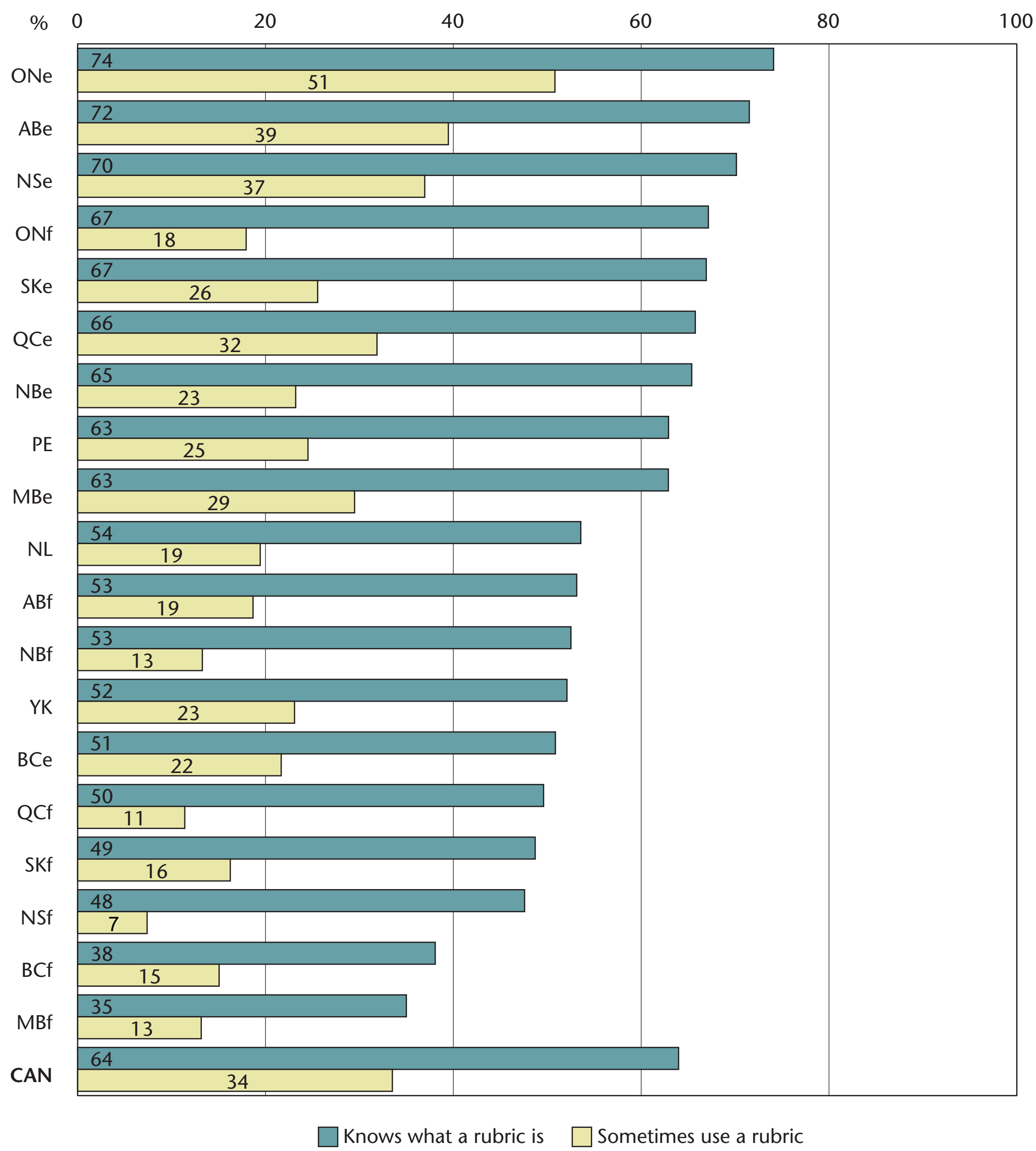


Chart 9-6 gives the results for frequency of use of rubrics for scoring. For Canada as a whole, more than half the students reported that rubrics are sometimes or often used for scoring. Use is highest in Ontario English and French and Nova Scotia English. Beyond this, there is little variation in the "often" category but more in the "rarely or never" category.

Chart 9-7 gives mean mathematics scores for students reporting knowledge of rubrics and use of rubrics in mathematics classes. Knowing what a rubric is and using rubrics at the start of assignments are both significantly positively associated with mathematics performance. Frequency of use for scoring shows a non-linear pattern but with students who report the most frequent use having the highest mean mathematics scores.

Chart 9-6 **Student reports of frequency of use of rubrics for scoring by jurisdiction and language**

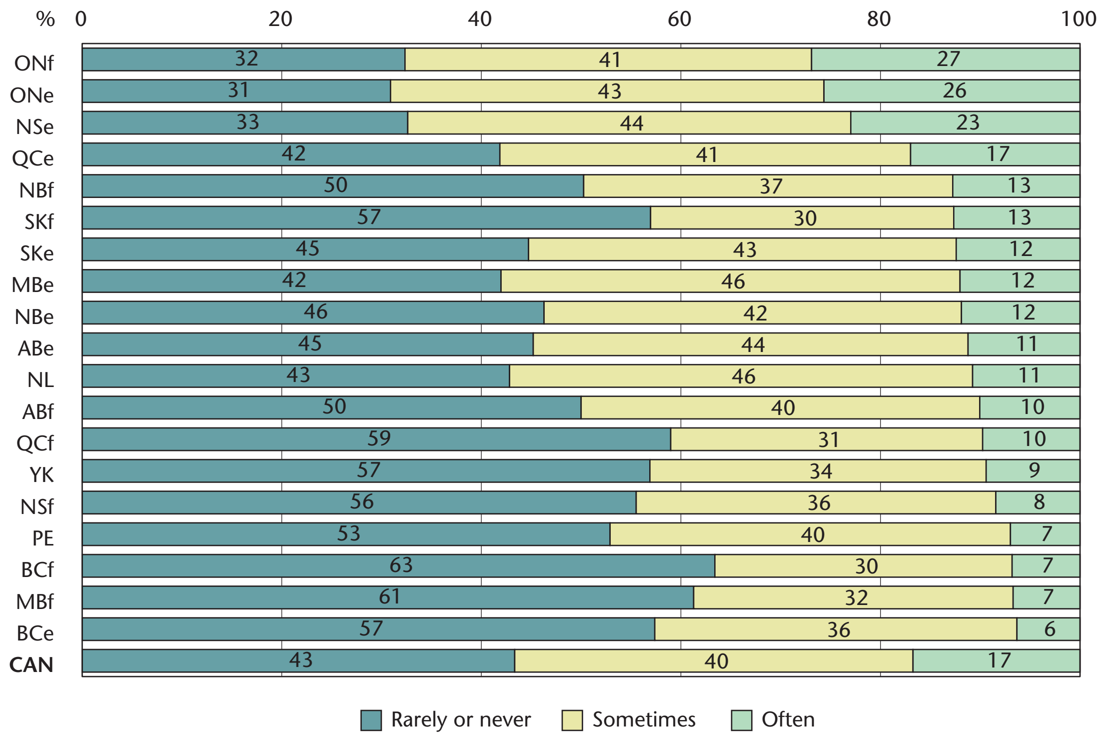
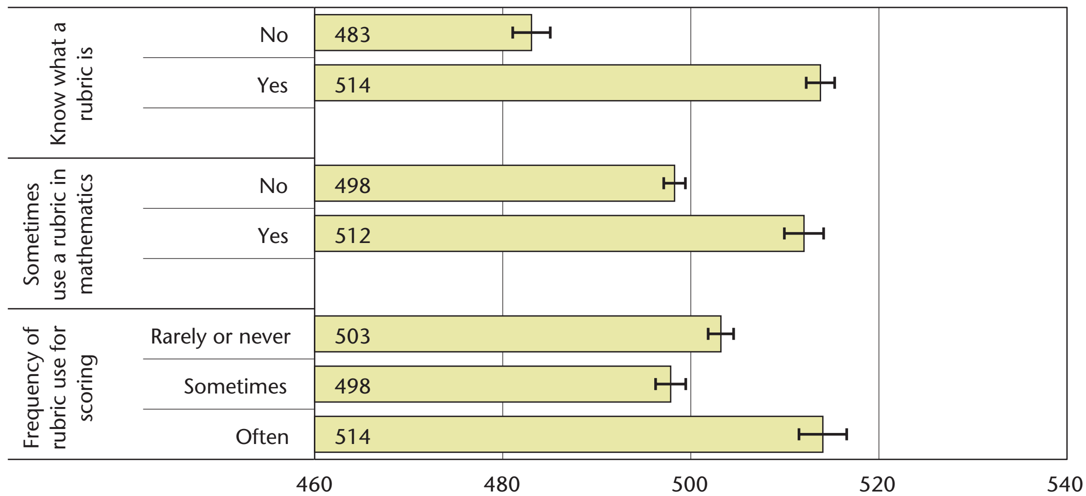


Chart 9-7 **Mean mathematics scores by knowledge and use of rubrics**



Types of assessment items used by teachers

Teachers were asked how frequently (rarely or never, sometimes, often) they use each of four different types of items or questions in assessing their students: selected response, short answer, extended response multi-step, and extended response with explanations. Responses by population are given in Chart 9-8. These are presented in descending order of use of "extended response multi-step" question, as this is the type most frequently used overall.

These results are fairly complex and are best examined within populations or jurisdictions. However, the following highlights are worth noting:

- Overall, francophone teachers use extended response items more frequently than do anglophone teachers.
- In most populations, there is wide variation in use of the different types, suggesting that there is a trade-off among the types. A notable exception is Newfoundland and Labrador, where teachers appear to use all types relatively frequently.

Chart 9-9 gives mean mathematics scores by frequency of use of each of these item types. The general pattern is that greater use of shorter item types (i.e., selected response and short response) is associated with lower mathematics performance, though the trend is not strong or linear. The pattern is clearer for the two kinds of extended response items, where greater use of such items is associated with higher scores. (The wide error bar for "rarely or never" use of extended response multi-step items occurs because few teachers responded in this category).

Chart 9-8 **Percentage of teachers "often" using item types by jurisdiction and language**

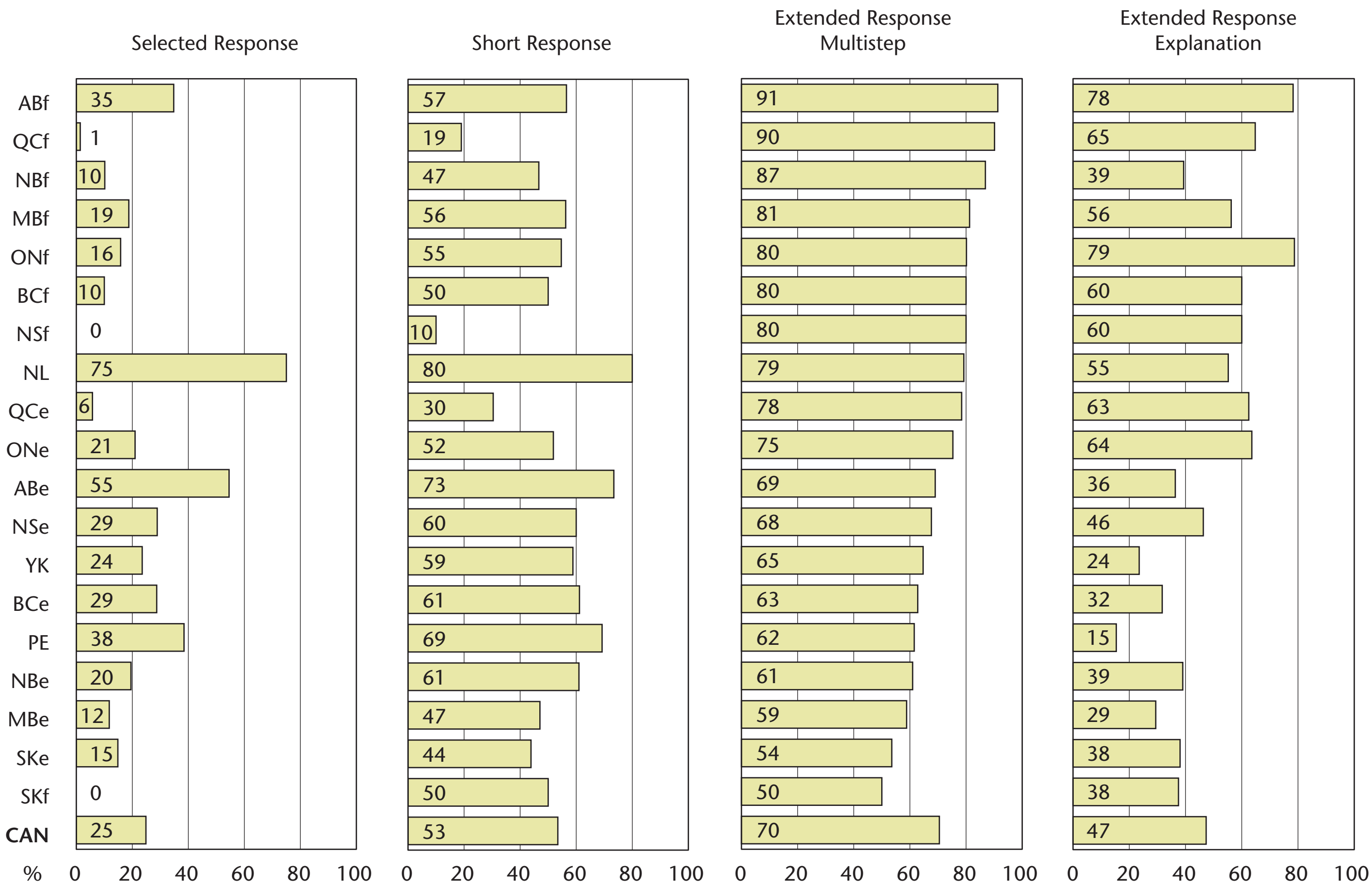
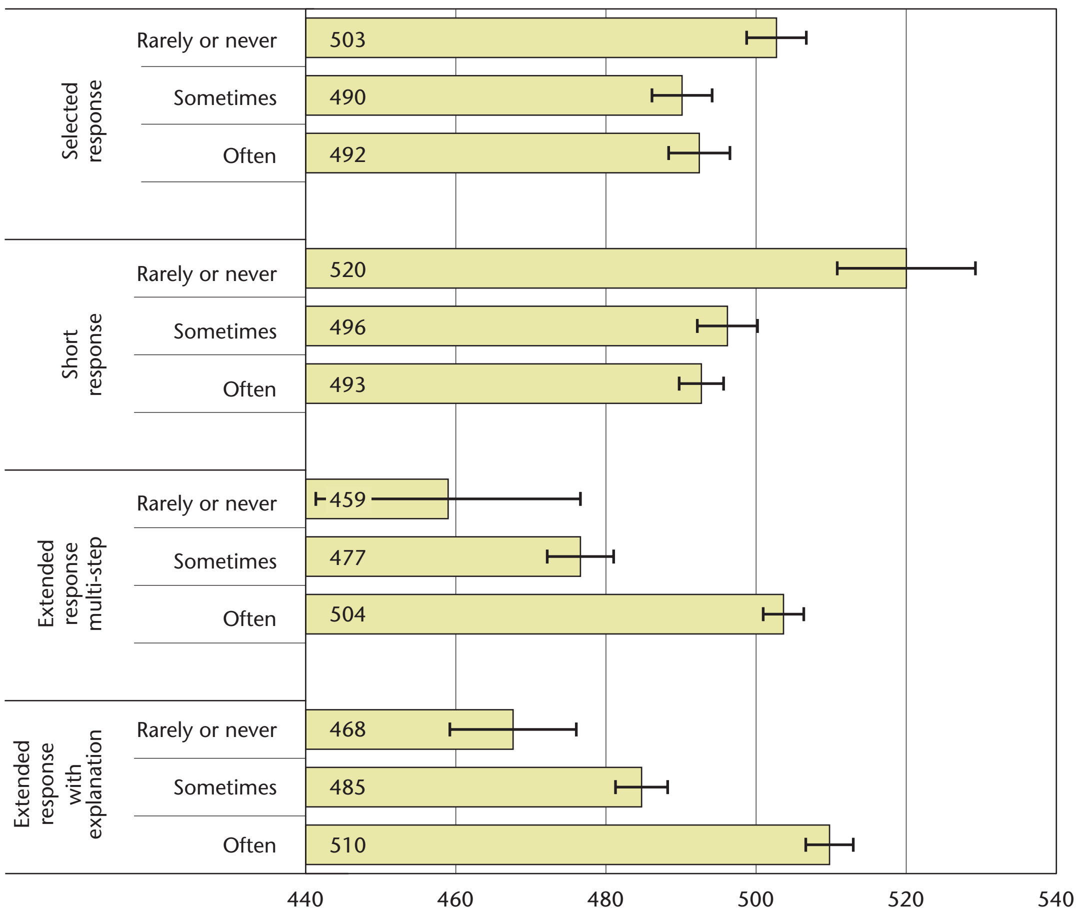


Chart 9-9 **Mean mathematics scores by item types**



Non-academic criteria in grading

Non-academic criteria that may be used for assessment purposes include attendance, class participation, improvement, effort, and behaviour. In each case, teachers were simply asked whether they assign marks on the basis of these elements. "Yes" responses to these items were summed to yield a scale from 0 to 5 on the number of these elements used. These were then combined into three categories for simplicity in reporting.

Chart 9-10 presents the percentages of teachers using 0–1, 2–3, and 4–5 of these elements for assigning marks. Overall, about two-thirds of teachers reported using none or one of these criteria. However, there are wide variations across populations. Alberta teachers stand out as using the fewest of these criteria and Saskatchewan French teachers the most.

Mean mathematics scores by number of these non-academic criteria used for grading are given in Chart 9-11. The pattern here is one of reduced mathematics performance with use of more of these criteria. However, only the difference between the lowest and the highest use is statistically significant.

Chart 9-10 **Number of non-academic criteria used to assign grades by jurisdiction and language**

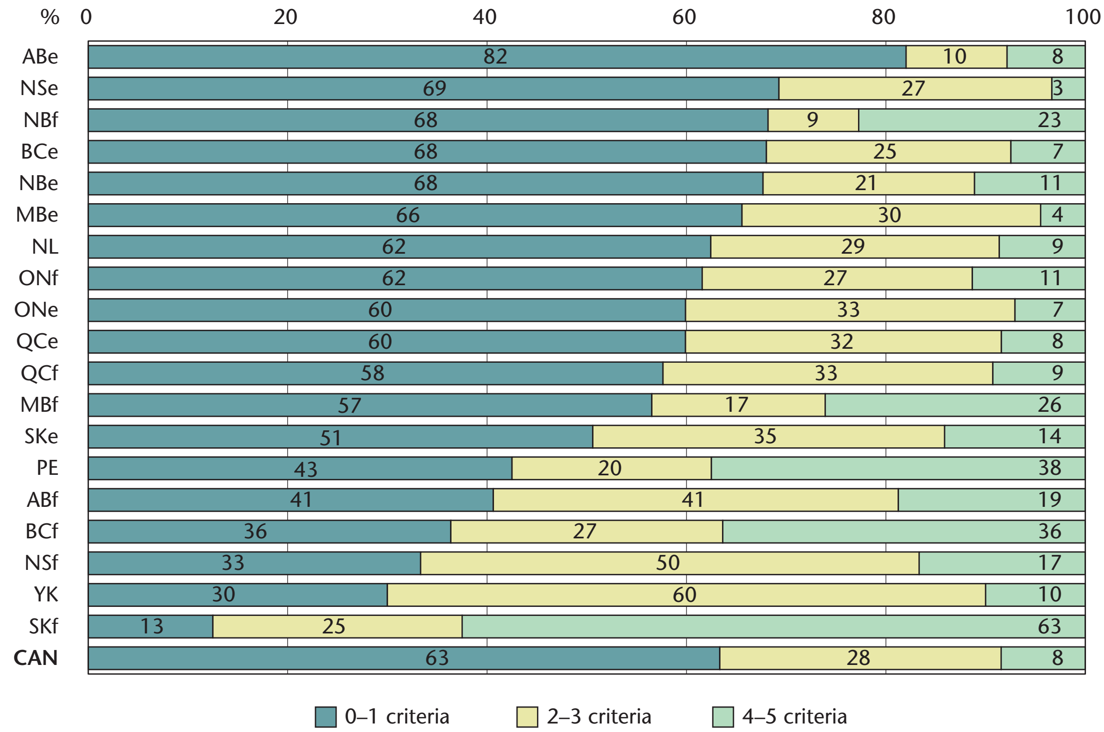
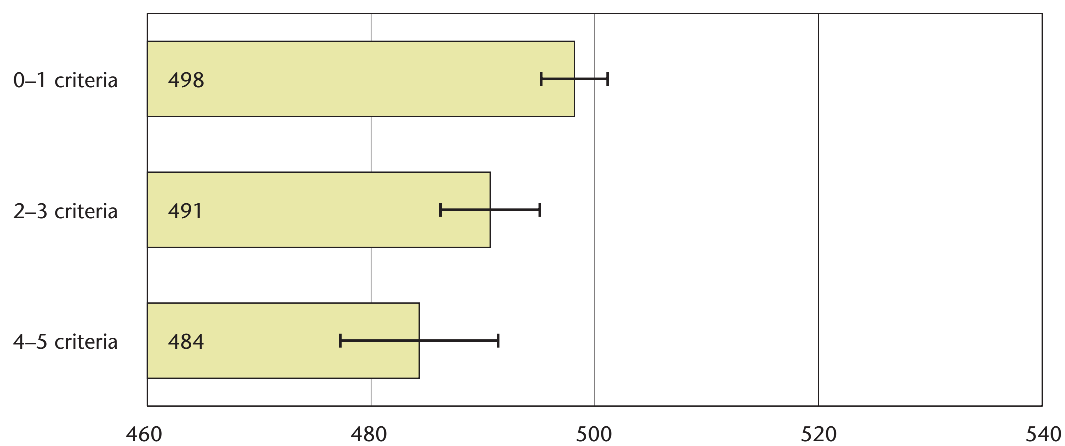
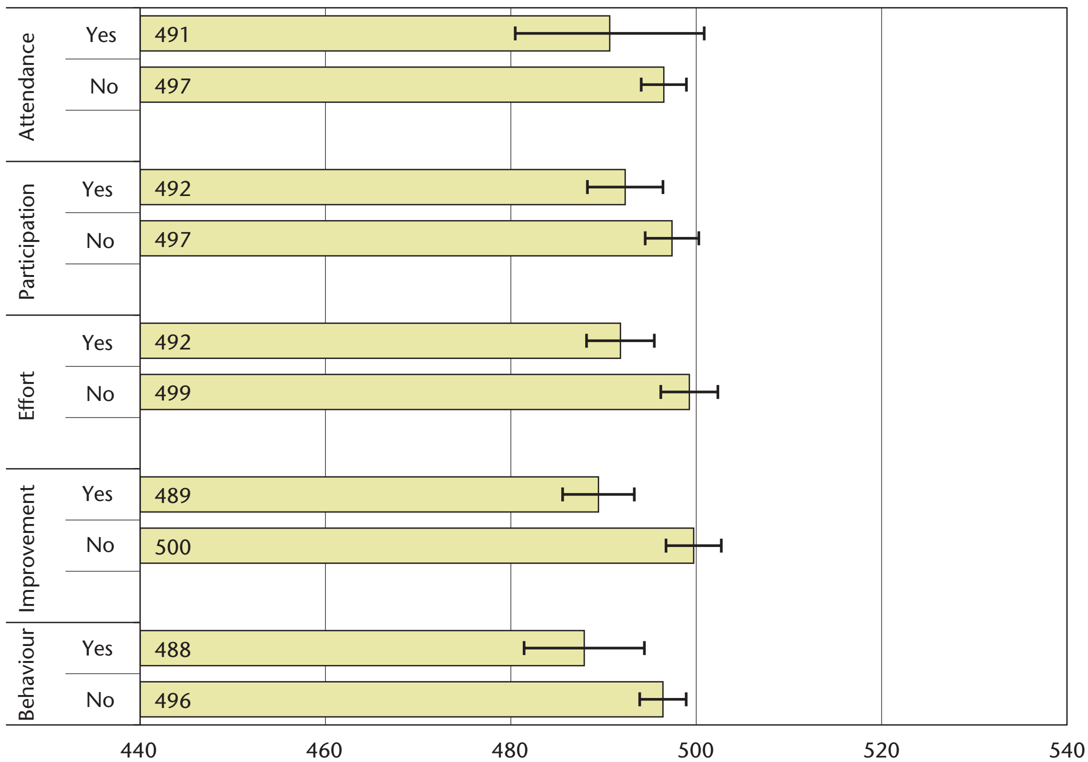


Chart 9-11 **Mean mathematics scores by number of non-academic criteria**



An analysis of the separate criteria is shown in Chart 9-12. The pattern is toward lower mathematics scores for teachers who use these non-academic criteria. However, the difference is statistically significant only for effort and improvement.

Chart 9-12 **Mean mathematics scores by specific non-academic criteria**



Assessment components contributing to student final marks

Questions in this area had to do with teacher use of eight different forms of assessment that contribute to students' final marks. Teachers were asked how often they use these forms of assessment for that purpose, using a three-point frequency scale (rarely or never, sometimes, often). The eight forms are:

- common school-wide tests or assessments
- teacher-made classroom tests
- assignments/projects
- homework
- portfolios
- self-assessment
- group work
- peer assessment

Factor analysis of this scale yielded a complex, difficult to interpret, factor pattern. Also, four of the types (portfolios, self-assessment, group work, and peer assessment) were used infrequently by teachers in all populations. It was therefore decided in this case to treat each of the remaining components separately.

Chart 9-13 shows the four most "often" used assessment types by population. Teacher-developed classroom tests are by far the most widely used, with relatively little difference across populations. Assignments and projects is the next most widely used, at 30 per cent nationally but with wider variation, from a high of 90 per cent in Yukon to a low of 3 per cent in Quebec French. Homework use also varies widely, from the 75 per cent to 80 per cent range (Yukon and Saskatchewan French) to less than 10 per cent (New Brunswick French and Ontario French). Use of common school-wide tests is fairly low overall and in most populations, with the notable exception of Newfoundland and Labrador, where 42 per cent of teachers reported that they often use such tests.

Chart 9-14 gives mean mathematics scores for these four assessment methods. School-wide tests or assessments and homework show no significant effects. For teacher-made classroom tests, the difference between "sometimes" and "often" is significant in favour of greater use. (The result for "rarely or never" is not particularly meaningful because few teachers are in this category, as indicated by the wide error bar.) The results for assignments/projects indicate that greater use of this method is associated with lower mathematics scores.

Chart 9-13 **Percentage of teachers "often" using selected assessment methods to assign grades by jurisdiction and language**

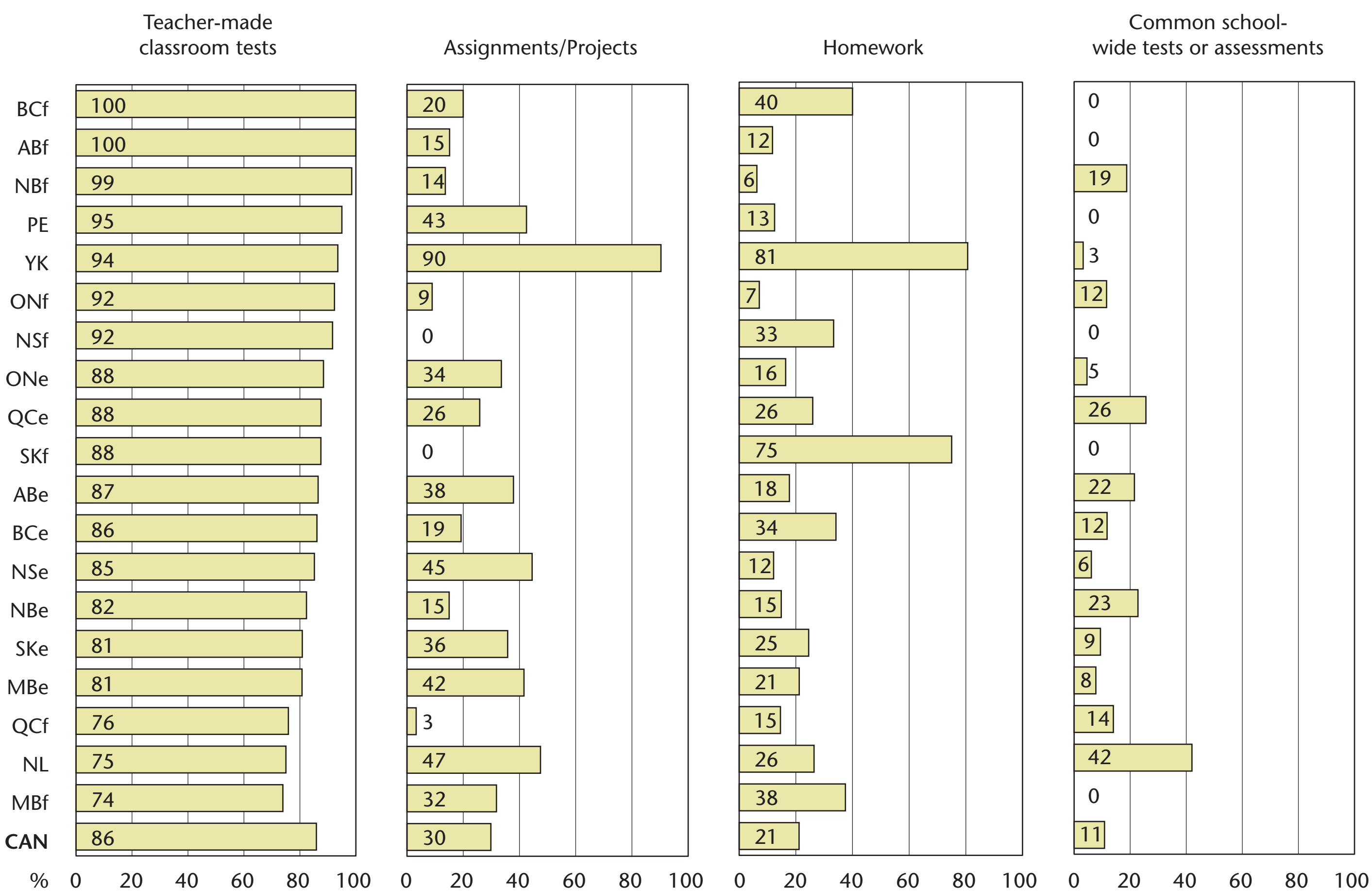
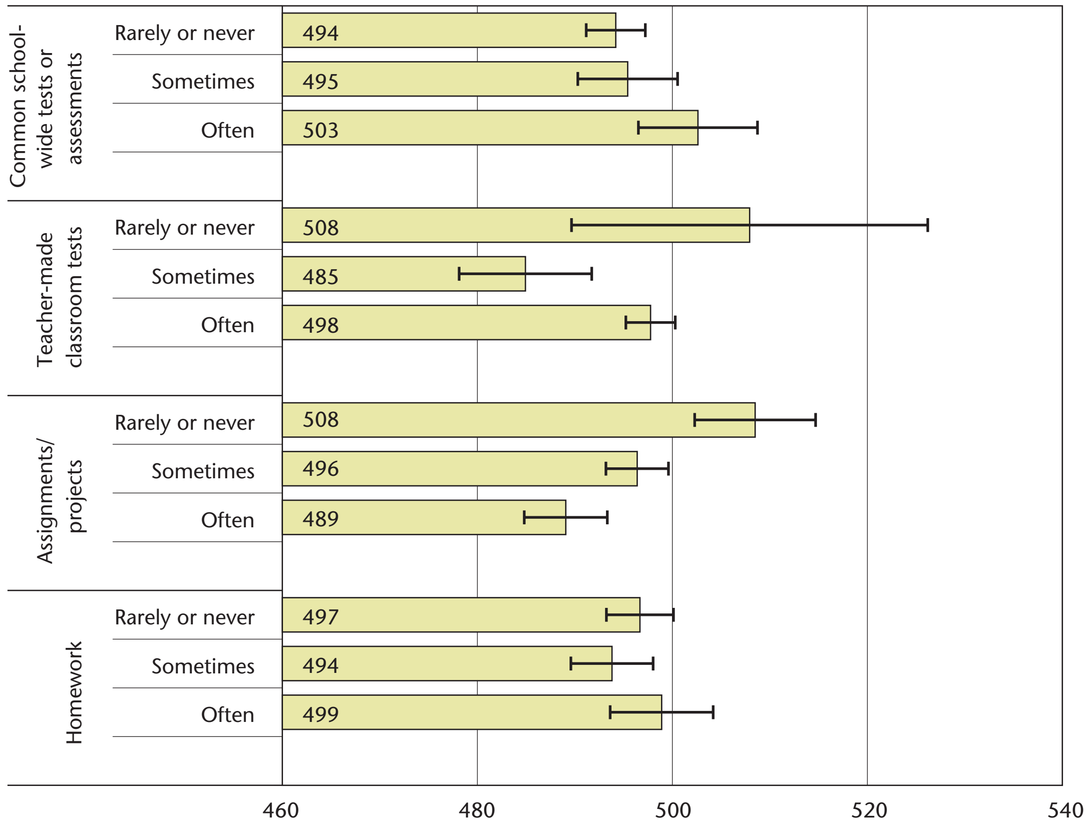


CHART 9-14 **Mean mathematics scores by assessment methods**



Grading methods

Teachers were asked if they used each of seven methods of final reporting, using a "yes"/ "no" scale, with multiple responses being allowed. The methods are:

- numeric grades
- comments
- descriptive levels
- letter grades
- numeric levels
- checklists based on course outcomes
- other

Chart 9-15 shows the four most commonly methods used by population. Numerical grades are used by more than 70 per cent of teachers in most populations. The exceptions are New Brunswick English and Nova Scotia English and French, where the percentages are much lower. The next most frequently used method is comments. However, there is much wider variation in use of this method, ranging from 88 per cent of Saskatchewan French teachers to 13 per cent of Yukon teachers. The remaining two methods, descriptive levels and letter grades, are used somewhat less often overall, again with wide variations across populations.

Chart 9-16 gives the mean mathematics scores for use of each of the methods (including those with less frequent use than those reported above). There are generally no significant differences between these methods. The exception is letter grades, which shows a significantly lower mean score than either numerical grades or comments.

Chart 9-15 **Methods of final reporting by jurisdiction and language**

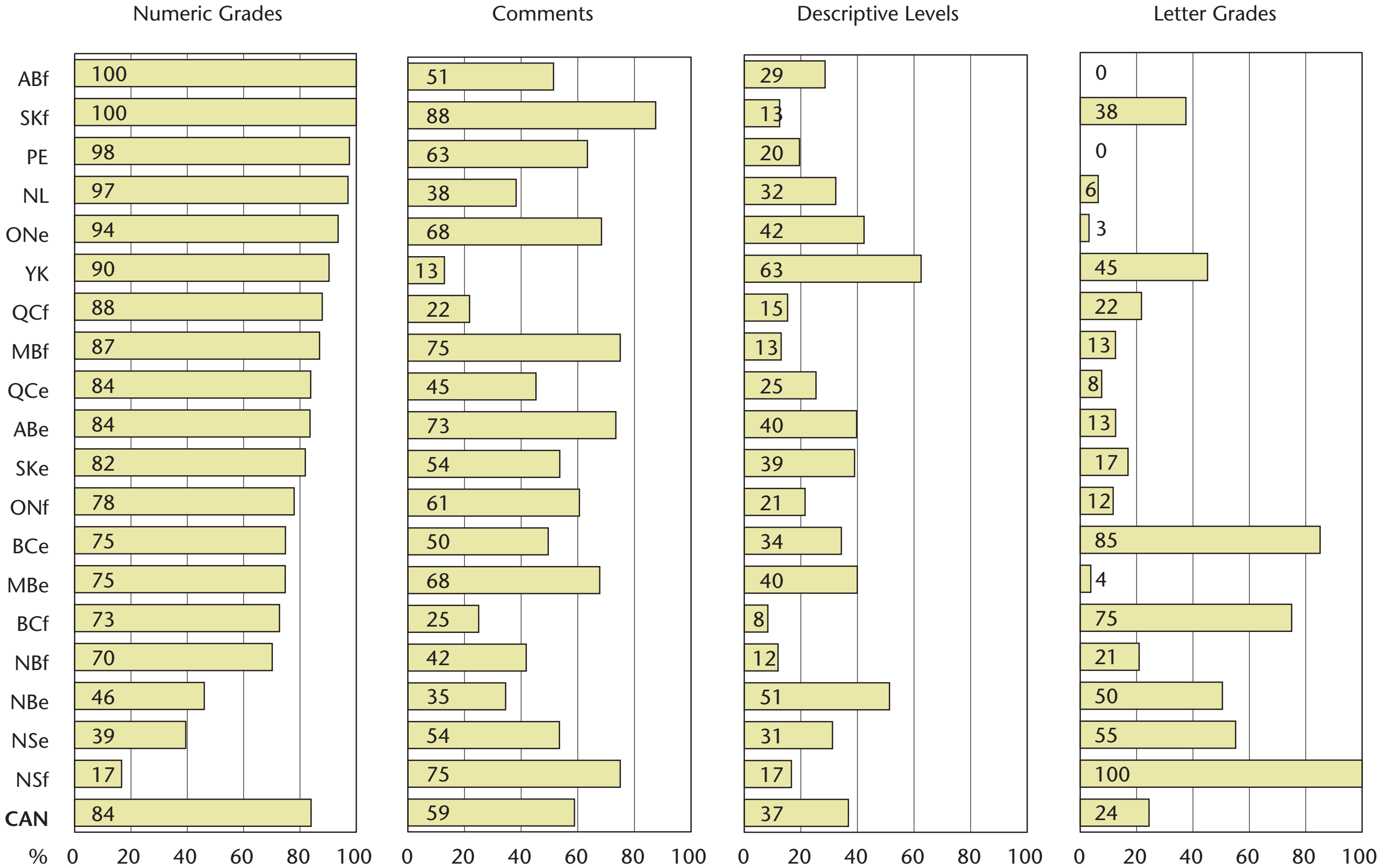
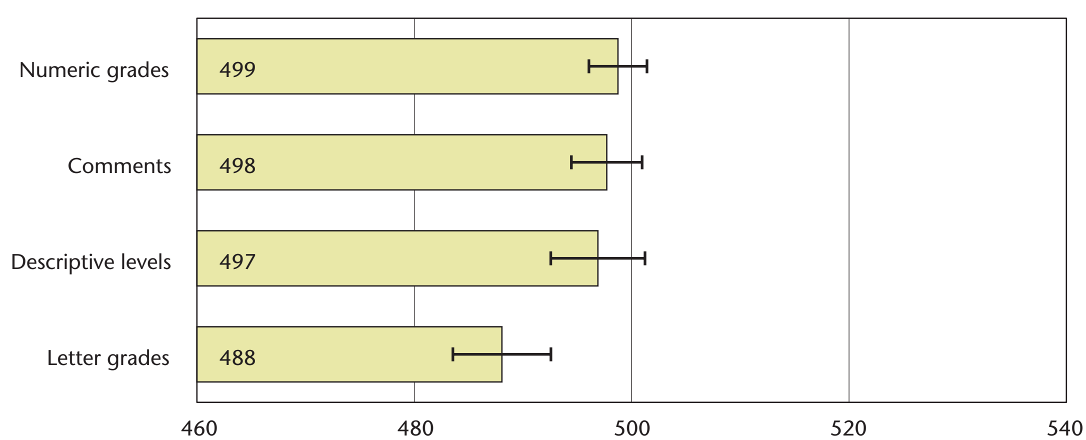


Chart 9-16 **Mean mathematics scores by methods of final reporting**



Availability and use of external assessments

Principals were asked to give their opinions on the availability and use of external assessments such as PCAP and PISA, using seven items on a 4-point Likert scale (strongly disagree, disagree, agree, strongly agree). These items yielded two factors as shown in Table 9-2. These results may be interpreted as associated with use of the results in the school or their more general availability outside the school, and ability to interpret these results. For the availability factor, negative factor loadings indicate that scores in the highest quartile (A) are to be interpreted as negative views on external assessments.

Table 9-2 **Principals' views on external assessments**

Factor	To what extent do you agree with the following statements about such assessments (PISA, PCAP)?
Availability of external assessments (–) results	• These test results are easily interpreted. • These test results are easily obtained. • These test results are easy to use in making instructional change.
Use of external assessments results	• In our school, we discuss these test results with groups of teachers or at staff meetings. • We discuss these results with parents/guardians in our school. • Teachers actually do use these test results to make changes in their instruction.
Note: (–) indicates factors that showed negative loadings of their items during the factor analysis.	

Chart 9-17 gives the mean factor scores by population for availability of results of external assessments such as PISA and PCAP. Only Saskatchewan English and French are above the Canadian average on this measure, indicating that principals in these populations have stronger negative views on availability than those in other populations. Eight populations, from Nova Scotia English to New Brunswick French (Manitoba French excluded because of the large error) on the chart, are below the Canadian average.

Results for use of external assessments are shown in Chart 9-18. Schools in Newfoundland and Labrador, New Brunswick English and French, and Nova Scotia English have more positive views on use of these assessments than the Canadian average. Schools in six populations, from British Columbia French to Quebec French, have less positive views on the use of external assessments than the Canadian average. Quebec French stands out as being lower on this measure than all others except Yukon.

Chart 9-17 **Mean factor scores for availability of external assessment results by jurisdiction and language**

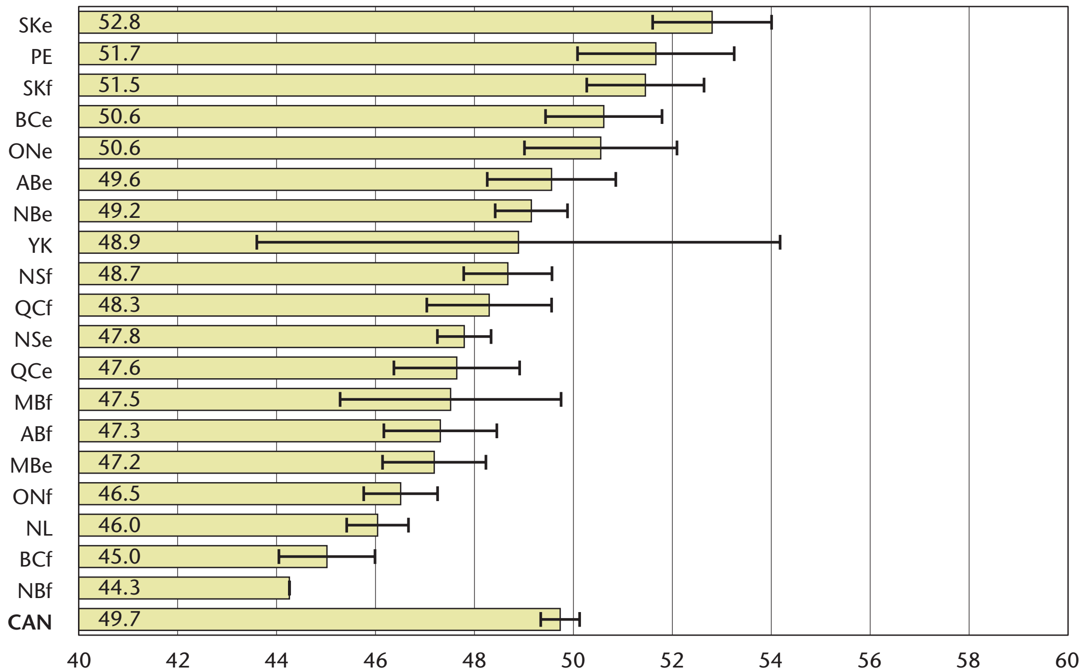


Chart 9-18 **Mean factor scores for use of external assessment results by jurisdiction and language**

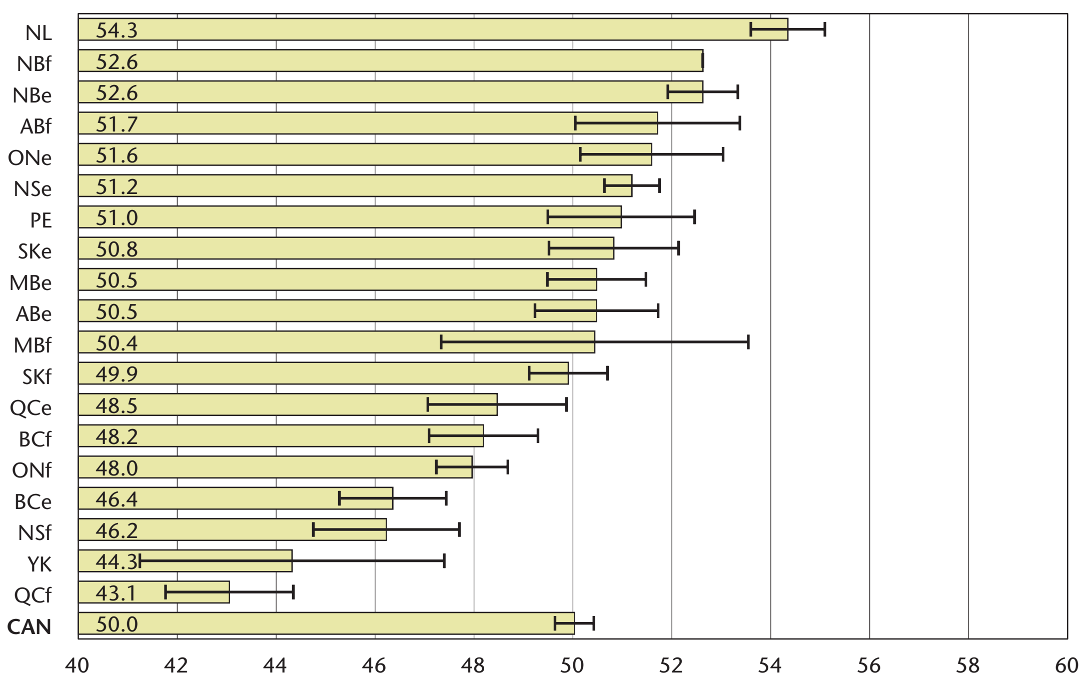
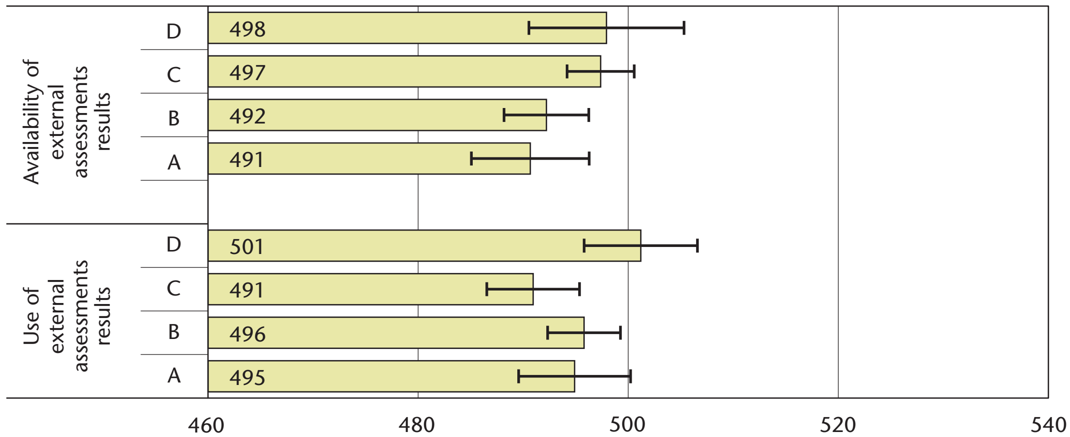


Chart 9-19 gives the mean mathematics scores for schools on availability and use of external assessments. No significant pattern of effects is evident for these variables.

Chart 9-19 **Mean mathematics scores by principals' views on the availability and use of external assessments**



A similar set of questions was asked about provincial/territorial assessments. This set was slightly different in that availability was not considered an issue. This set also gave two factors, with a slight difference in interpretation from the previous set, as shown in Table 9-3. In this case, the first factor reflects positive views on the use of these assessments. The second factor is interpreted as conveying negative attitudes toward provincial/territorial assessments.

Table 9-3 **Principal views on provincial/territorial assessments**

Factor	To what extent do you agree with the following statements about provincial/territorial assessments?
Use of provincial/territorial assessment results	• In our school, we discuss these test results with groups of teachers or at staff meetings. • We discuss these results with parents/guardians in our school. • Teachers actually do use these test results to make changes in their instruction. • These test results are easy to use in making instructional changes. • These test results are easily obtained. • Principals have a responsibility to develop an action plan in response to these results. • These test results are easily interpreted.
Negative attitude toward provincial/ territorial assessment results	• School-level results from these tests should be published in newspapers (–). • These tests take too much time away from teaching and learning.
Note: (–) indicates items that showed negative loadings during the factor analysis.	

Charts 9-20 and 9-21 give the mean factor scores for each of these variables by population. On the use factor, seven populations, from Newfoundland and Labrador to Nova Scotia English, are above the Canadian average, and seven others, from Yukon to Quebec French, are below.

Negative attitudes toward provincial/territorial assessments are strongest among Saskatchewan English principals. Four other populations, from Alberta French to Manitoba English, also show negative attitudes stronger than the Canadian average. Six populations, from Nova Scotia English to Quebec French on the chart (again excepting Yukon because of the large error) have views less negative than the Canadian average.

Chart 9-20 **Mean factor scores for use of provincial/territorial assessment results by jurisdiction and language**

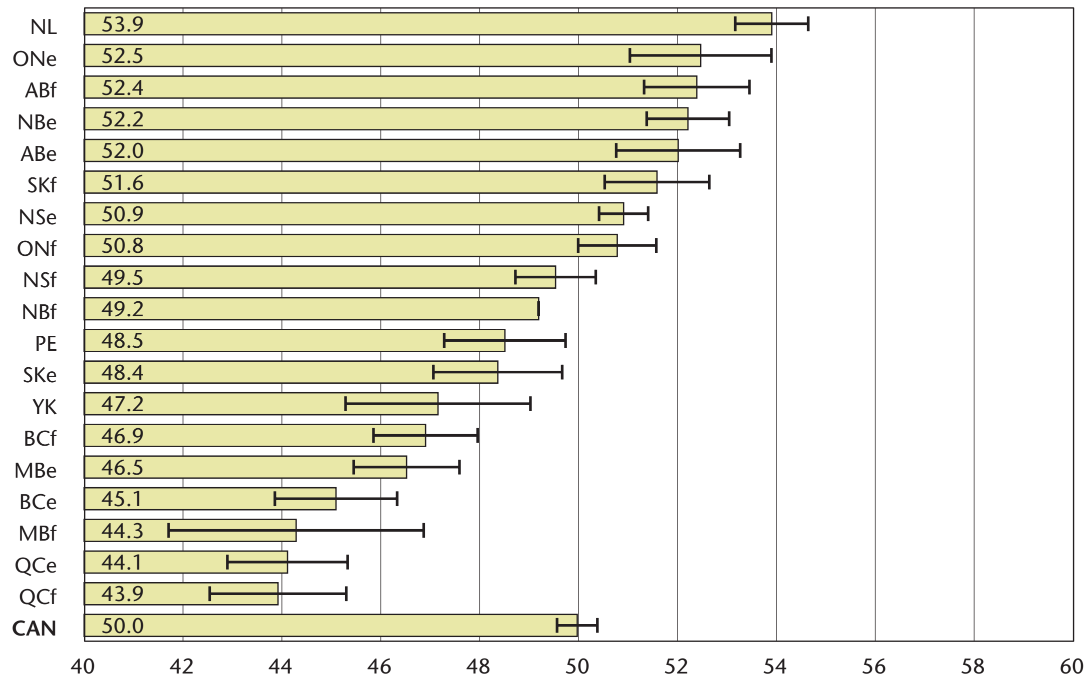


Chart 9-21 **Mean factor scores for negative attitudes toward provincial/territorial assessment results by jurisdiction and language**

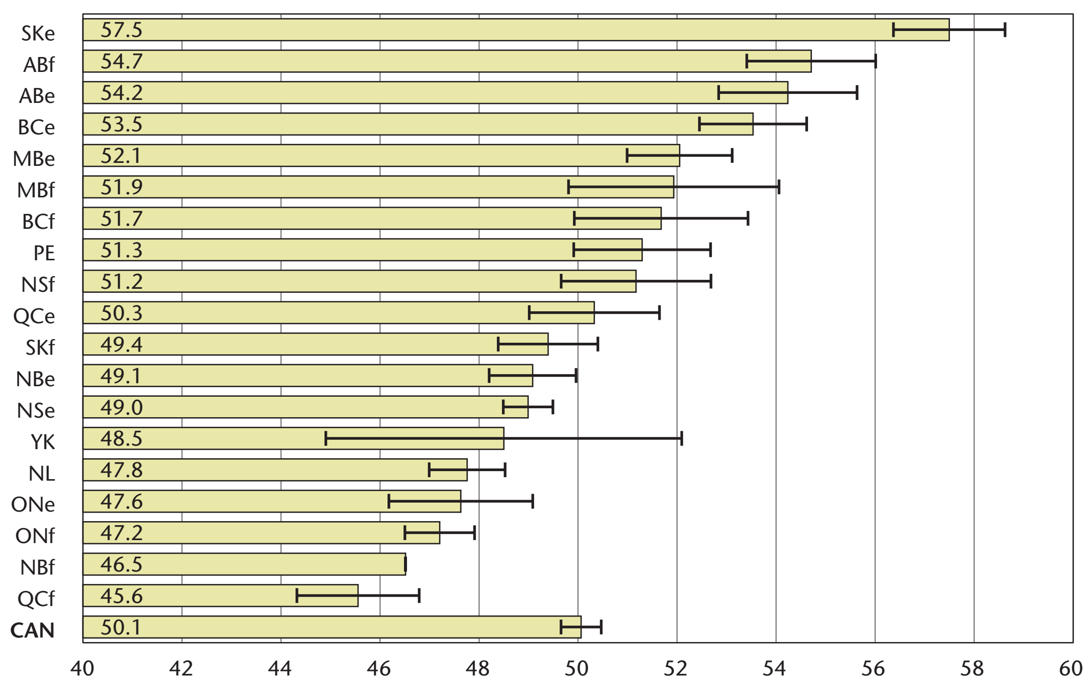
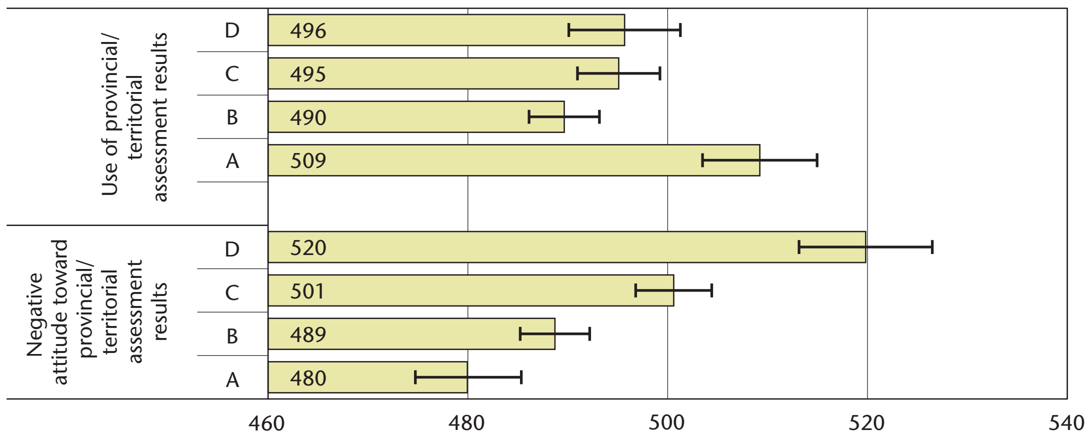


Chart 9-22 gives the mean mathematics scores for these two variables. The use variable shows a significant effect only for those with the most highly positive views on the use of provincial/territorial assessments. The pattern for negative attitudes is more distinctly linear. Schools whose principals have less negative attitudes have higher mathematics scores.

Chart 9-22 **Mean mathematics scores by principals' views on provincial/territorial assessments**



Purpose for which assessment results are used

Principals were asked to report on the frequency of use (rarely or never, sometimes, often) of three types of assessments, classroom assessments, provincial/territorial assessments, and pan-Canadian or international assessments for a variety of purposes, including grading and reporting on individual student progress, program evaluation, and teacher effectiveness. To reduce the complexity in reporting these results, a composite frequency scale was developed for each type of assessment by summing the three response categories for each type across all of the uses. These summed scores were then regrouped to yield a composite three-point scale for each assessment type corresponding to the original scale.

Charts 9-23 to 9-25 give the results for these variables, with populations ordered by the most frequent response category. As might be expected, classroom assessments are used more often overall than the other forms. There is substantial variation across populations in the extent of use of both classroom and provincial/territorial assessments. National assessments are rarely used within most populations.

Chart 9-23 **Use of classroom assessments for various purposes by jurisdiction and language**

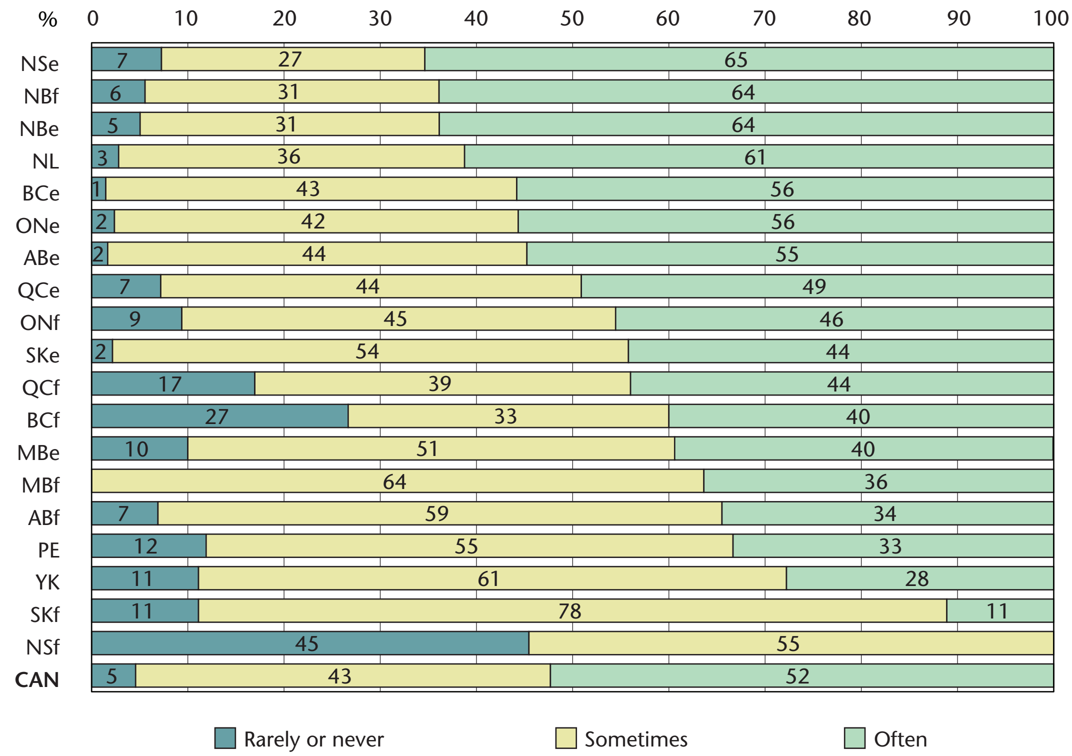


Chart 9-24 **Use of provincial/territorial assessments for various purposes by jurisdiction and language**

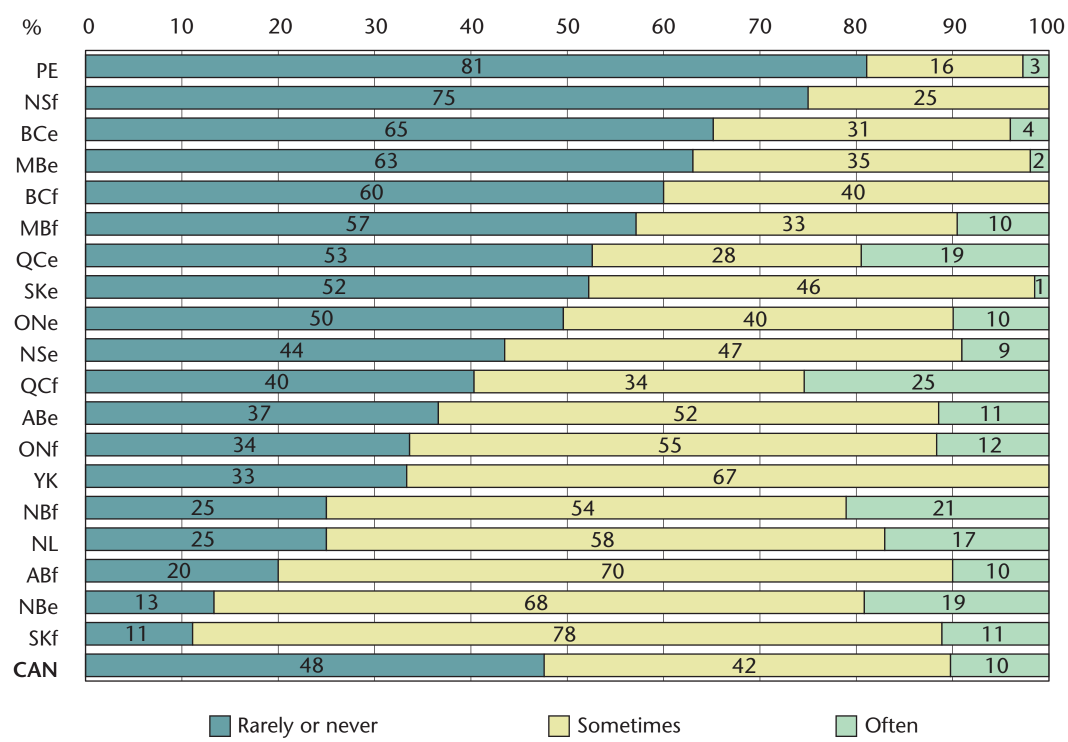
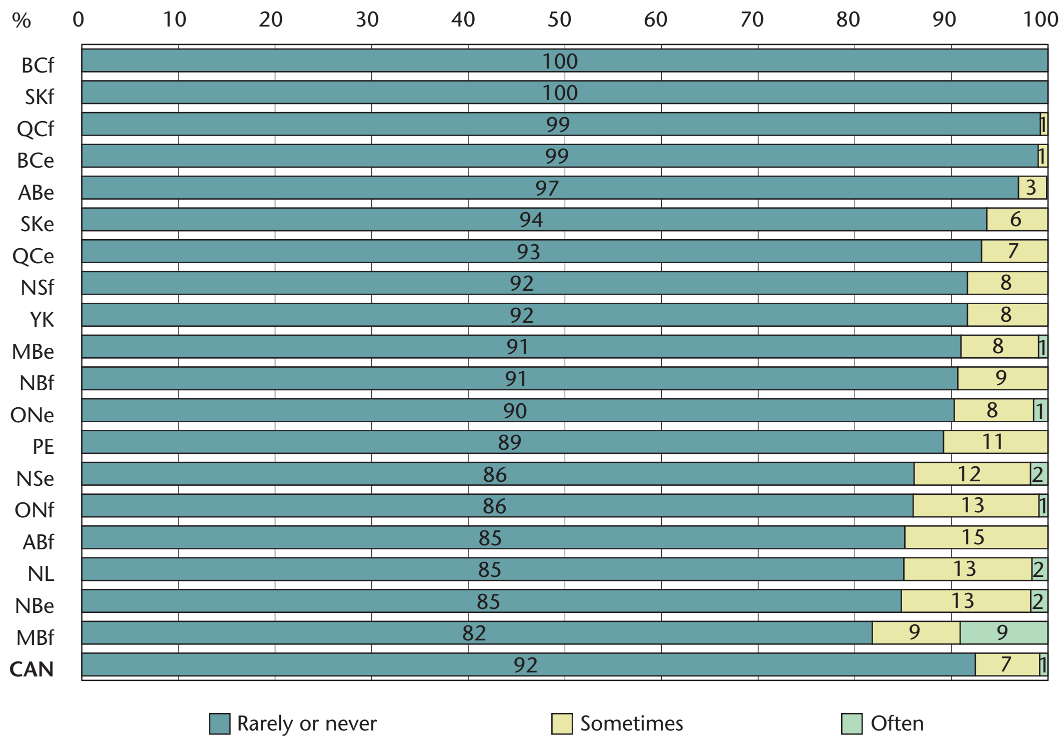


Chart 9-25 **Use of national assessments for various purposes by jurisdiction and language**



The above scales are not particularly useful in examining the impact on achievement of using these forms of assessment. However, examining each type of use separately yields some interesting results. Because of the large number of comparisons possible with this set, and the small frequency of use in some cases, only those uses that show statistically significant effects are reported. These are given in Chart 9-26.

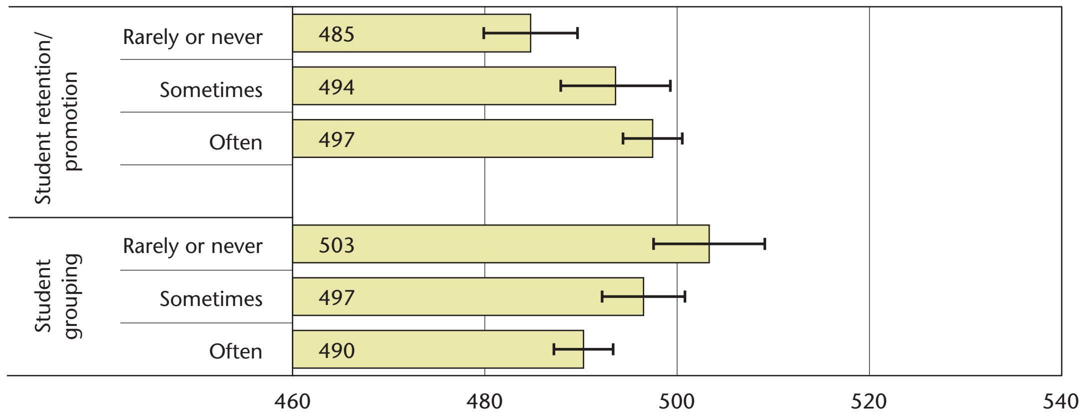
The classroom assessment results show a pattern of increasing mathematics scores with increased use of assessment for student retention and promotion. (Only the difference between the lowest and highest categories is statistically significant.) The opposite is true for use of classroom assessment for student grouping, for instruction.

Use of provincial/territorial assessment results to judge teacher effectiveness shows a non-linear pattern, with statistically significant differences between all categories, but the lowest score for the middle category. Overall, however, "often" using such assessments for this purpose yields the highest mathematics scores.

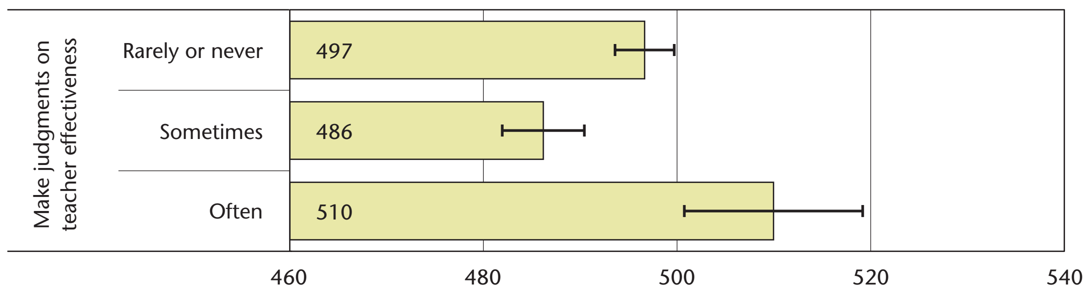
Use of pan-Canadian or international assessments shows a pattern of increasing mathematics scores with greater use of these assessments to inform parents of student progress and to monitor school programs, though not all differences are statistically significant. While it might seem reasonable to use such assessments for these purposes, in reality their results cannot usually be used in these ways because only samples of schools are used and school-level results are not reported. It is possible that these results reflect principals' attitudes toward such uses, in which case a positive view of these uses of large-scale assessments can be said to be positively associated with achievement. This is consistent with the earlier finding that a negative view is associated with lower achievement.

Chart 9-26 **Mean mathematics scores by selected uses of assessment types**

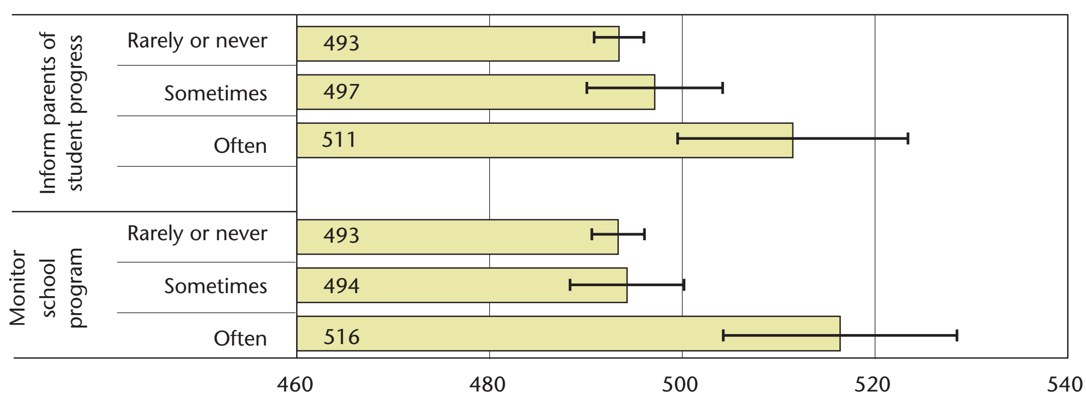
Classroom assessment



Provincial/territorial assessment



Pan-Canadian or international assessment



Multiple regression effects

Because of the large number of assessment variables, the modelling for this section was divided into three groupings, with student-, teacher-, and school-level variables entered as separate clusters.

Chart 9-27 gives the regression coefficients for the student-level assessment variables. Again, these coefficients should be interpreted as the change in mathematics score for one unit change in the predictor variable. In this case, the pattern is similar for the simple and multiple regression models, indicating that controlling for other variables in this set does not change the effect of any one variable. Using unconventional assessment (journals, portfolios, self-assessment, etc.) shows a significantly negative effect in both models. Using conventional assessment, knowing what a rubric is, and using a rubric when starting an assignment show significantly positive effects in both models.

CHART 9-27 **Regression coefficients for student assessment variables**

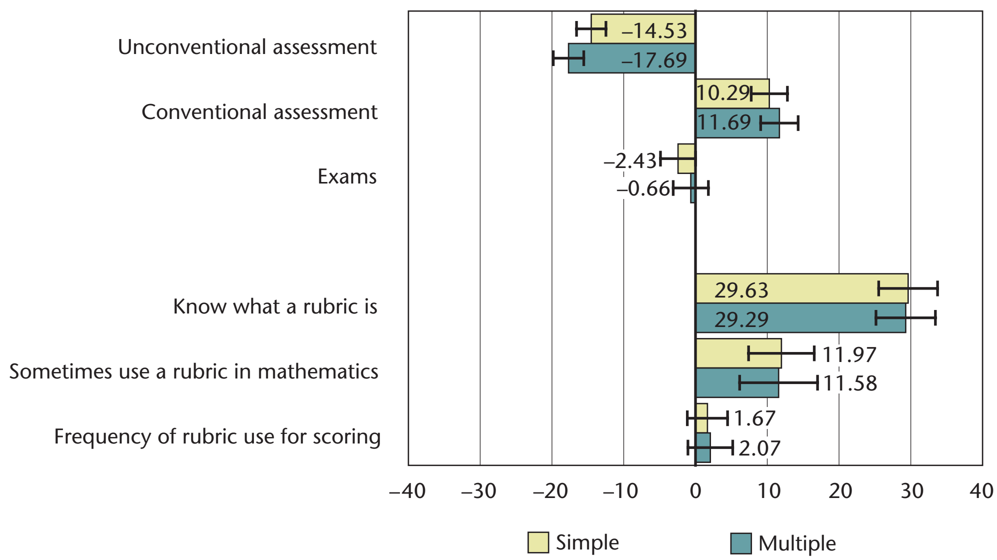


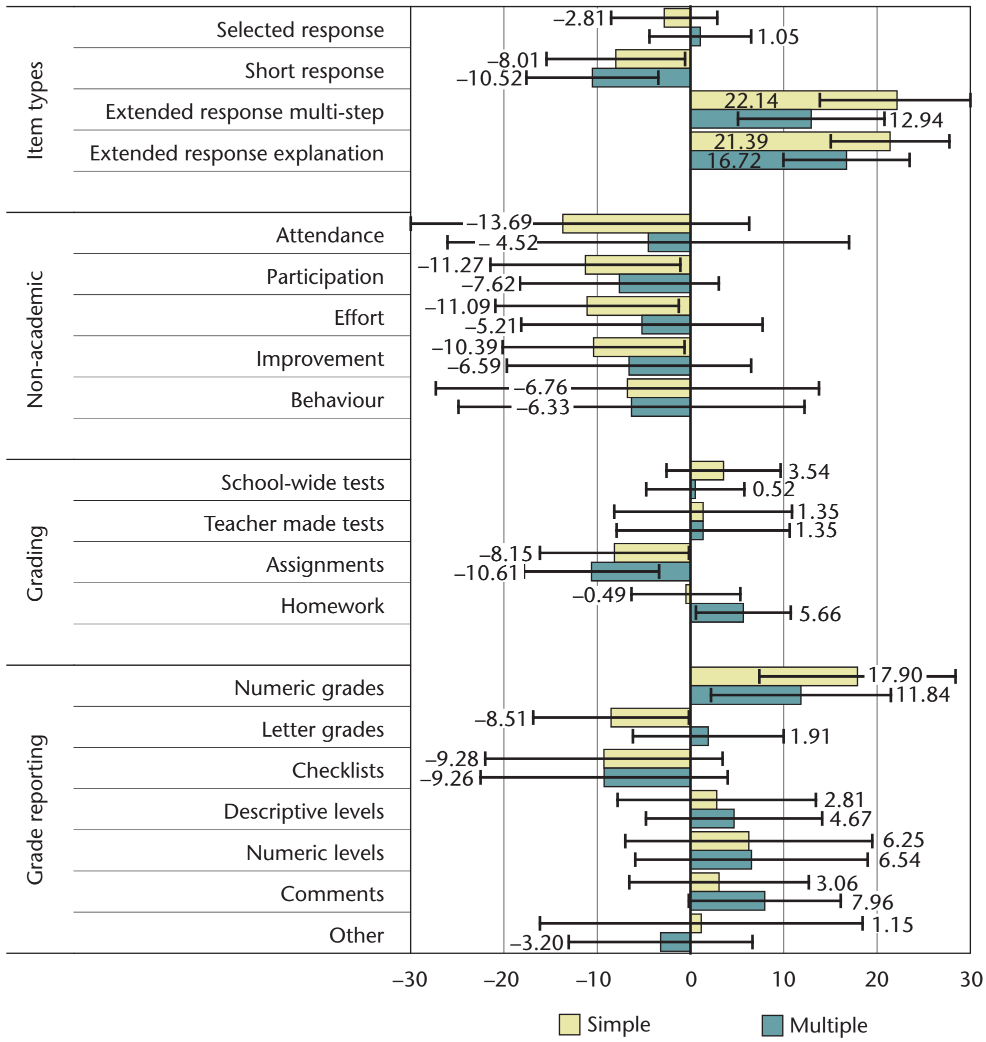
Chart 9-28 shows the coefficients for the teacher variables. The wide confidence intervals for many of these variables mean that relatively few statistically significant effects are found. The notable exception is the item type variable group, where using short response items has a significant negative effect, and using both forms of extended response items have significantly positive effects.

For the use of non-academic criteria, the effects for participation, effort, and improvement go from significantly negative in the simple regression model to non-significant in the multiple regression model, indicating that the simple regression effects are attenuated when other variables are controlled.

The use of assignments for grading is significantly negative in both models, while use of homework goes from non-significant in the simple regression model to significantly positive in the multiple regression model. The latter indicates that the homework effect is suppressed by other variables that are not accounted for in the simple regression model.

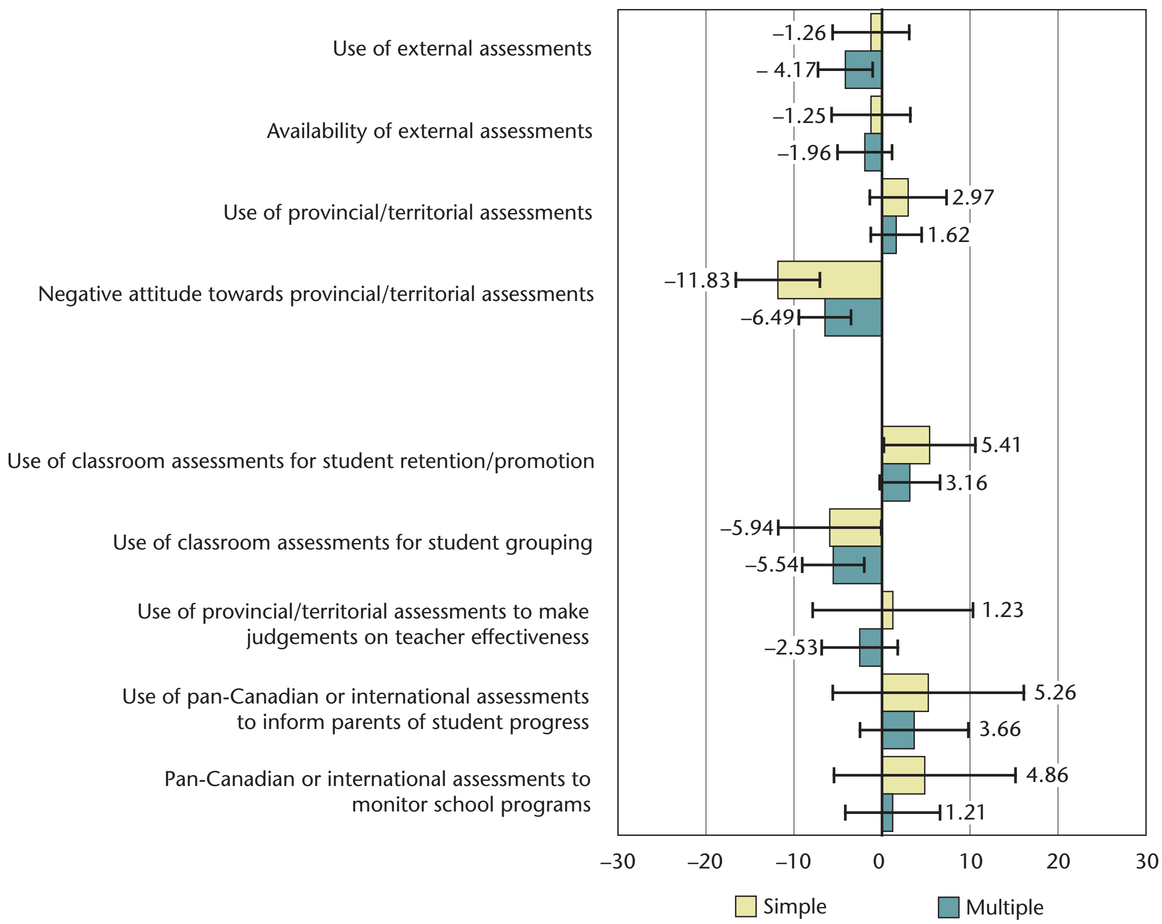
Finally, for methods of grade reporting, only the use of numeric grades has a significantly positive effect in both models.

CHART 9-28 **Regression coefficients for teacher assessment variables**



The coefficients for school assessment variables, given in Chart 9-29, show few statistically significant effects of these variables on mathematics achievement and mostly only small non-significant changes from the simple to the multiple regression model. Use of external assessments changes from non-significant to significantly negative in the multiple regression model. Holding negative views on external assessment is significantly negative in both models. The use of classroom assessment for student retention and promotion changes from significantly positive in the simple regression model to marginally non-significant in the multiple regression model. The opposite is true for use of classroom assessment for student grouping, which shifts from marginally negative to significantly negative. Both of these effects should therefore be treated as marginal.

Chart 9-29 **Regression coefficients for school assessment variables**



10 POPULATION DIFFERENCES AND ACHIEVEMENT EQUITY

The argument is commonly made that jurisdictions should strive not only for high average achievement but also for greater equity in achievement between its schools and its students. Indeed, "equality of opportunity" is a frequently stated goal of education systems, and is embodied in funding formulas, school programs, teacher allocations, and other policy instruments at the jurisdictional level. The assumption seems to be made that providing equality of opportunity will also help reduce disparities in outcomes between students, schools, and jurisdictions. The results of international studies, and more specifically the PISA results, indicate that Canada is one of the few countries to show both high achievement and relatively small differences between the lowest and the highest performing students.

The equity argument can be applied to differences between jurisdictions. The difference in PCAP 2010 mathematics performance between the highest and lowest performing jurisdictions is about 75 points or three-fourths of a standard deviation on the PCAP scale. Differences of this magnitude have been typical of PCAP and the earlier SAIP assessments. These differences have not changed much over time, and the relative rankings of jurisdictions have also been relatively stable.

It is reasonable to argue that, ideally, average scores for the lowest performing jurisdictions should come closer to those of the highest performing ones. Indeed, the scores of the highest performing jurisdictions might be seen as a benchmark toward which all jurisdictions should strive. One of the purposes of reports such as this one is to go beyond the simple reporting of population differences and to examine factors that contribute to such differences and that might help jurisdictions formulate policies to reduce these differences.

The goal of this chapter is to examine the equity issue as it relates to achievement in mathematics[21] at the student, school, and population (using both jurisdiction and language as in earlier chapters) levels.

Mathematics achievement variation within populations

PCAP achievement scores are scaled to a national weighted mean of 500 and a standard deviation of 100. The main focus of the *PCAP 2010 public report* is on differences in mean scores across populations. However, differences between schools and between students within populations are also of interest from an equity perspective. These differences are examined in two ways: first through the differences in "variance" across populations and second through differences between the highest and lowest scoring groups of students within each population, based on a division by standard deviation units, as given in Table 4-2 (p. 55).

Variance is defined as the square of the standard deviation and is used because it can be partitioned into student and school components in multi-level models of the type used in previous chapters. Thus variance can be used as an indirect measure of equity in that lower variance can be seen as higher equity. Chart 10-1 shows the total variance

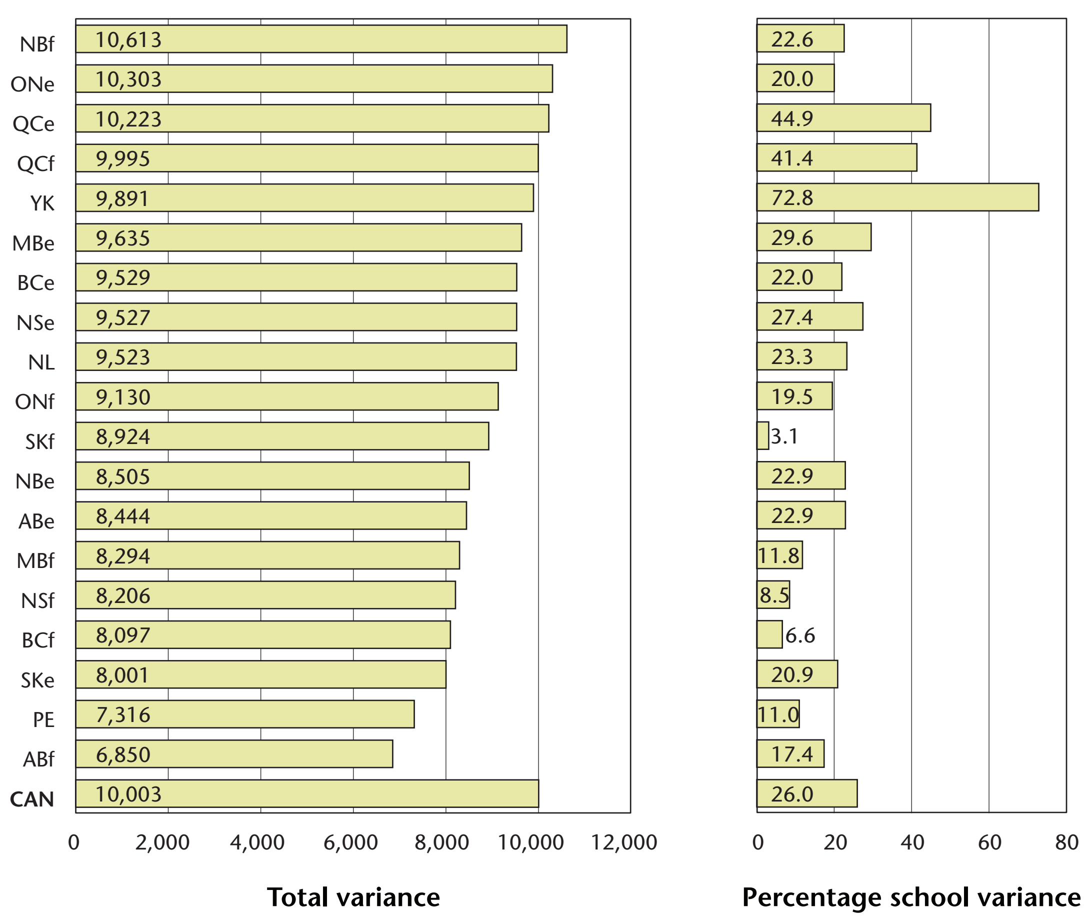
[21] A more comprehensive examination of the differences between jurisdictions and language groups and on factors contributing to these differences can be found in the following report: Council of Ministers of Education, Canada. (2012). *PCAP-13 2007: Jurisdictional profiles and achievement equity*. Toronto, ON: Author.

in mathematics scores by population and the proportion of that variance that can be attributed to differences between schools (the remainder represents differences between students within schools).

The first part of Chart 10-1 shows that about half the populations have total variance close to (within about 5 per cent) the Canadian average. Thus, none of the populations may be seen as showing much greater inequality among students than is typical for Canada. A few, including Saskatchewan English, Prince Edward Island, and Alberta French have total variance 20 per cent or more below the Canadian average. These populations may be seen as the ones who have achieved the greatest equity in student mathematics performance.

The picture across schools within populations is one of much greater variation. Yukon stands out as having close to three-fourths of its total variance being across schools and therefore, only about one-fourth of its variance across students. This suggests that, in Yukon, the student population within schools is relatively homogeneous compared to the differences between schools. The two Quebec populations also have relatively high between-school variation. On the other hand, a few small francophone populations, specifically, Saskatchewan French, Nova Scotia French, and British Columbia French, show very little variation across schools. Prince Edward Island also shows relatively low between-school variation. In all of these cases, the interpretation is that the schools in these populations are quite alike, no matter what the variation may be between students within schools.

Chart 10-1 **Total variance and percentage of variance across schools by jurisdiction and language**

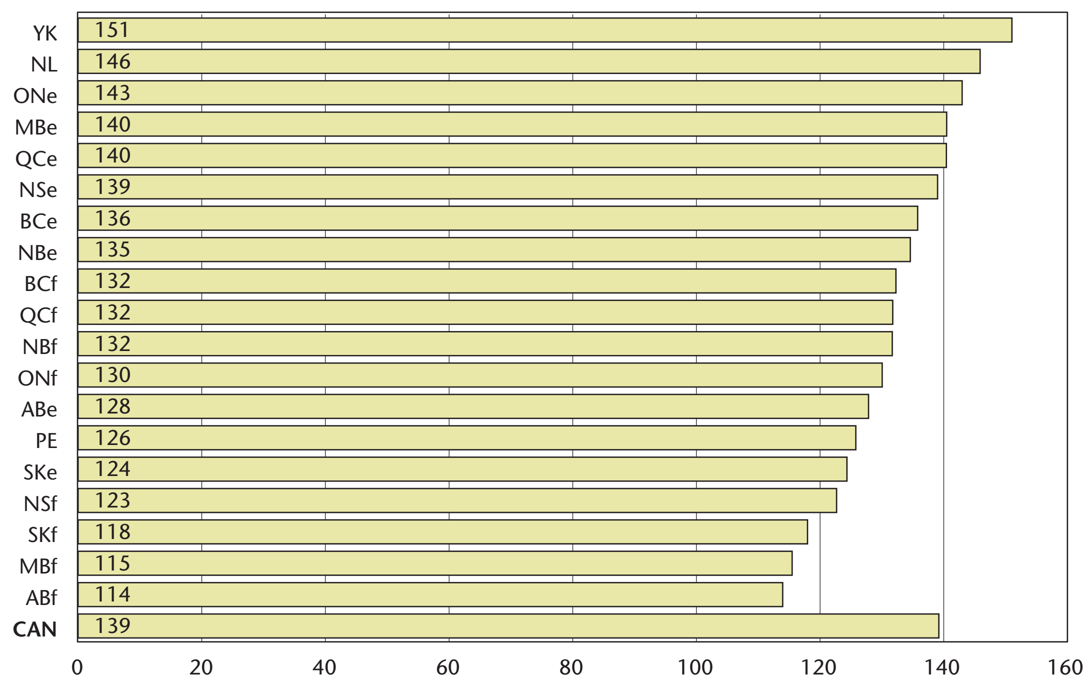


Mathematics interquartile range by population

Another way of looking at equity more directly in terms of scores is to divide the score distributions into quartiles (four groups of about equal size) and compute the cut score for each quartile. The "interquartile range," or the difference between the cut points for the 25th and 75th quartiles, may be used as an index of equity. This is a "reverse index," with higher numbers indicating less equity. The results for this approach are given in Chart 10-2.

This gives a somewhat different picture from that of Chart 10-1. The largest overall variation is found for Yukon, with Chart 10-1 indicating that this is largely a function of differences between schools. On the other hand, Newfoundland and Labrador and Ontario English, which have close to average overall variance, have relatively large interquartile ranges. Taken together, these two results suggest that these populations have a relatively large number in the two extreme quartiles while having relatively fewer at the furthest extremes beyond what is shown by the quartiles. This is because the extremes contribute a large amount to the total variance but less to the quartile breakdown. Four of the small francophone populations, Alberta French, Manitoba French, Saskatchewan French, and Nova Scotia French, show relatively small variation by the interquartile measure. These also tend to have relatively low overall variance and low between-school variance.

Chart 10-2 **Interquartile ranges for mathematics scores by jurisdiction and language**

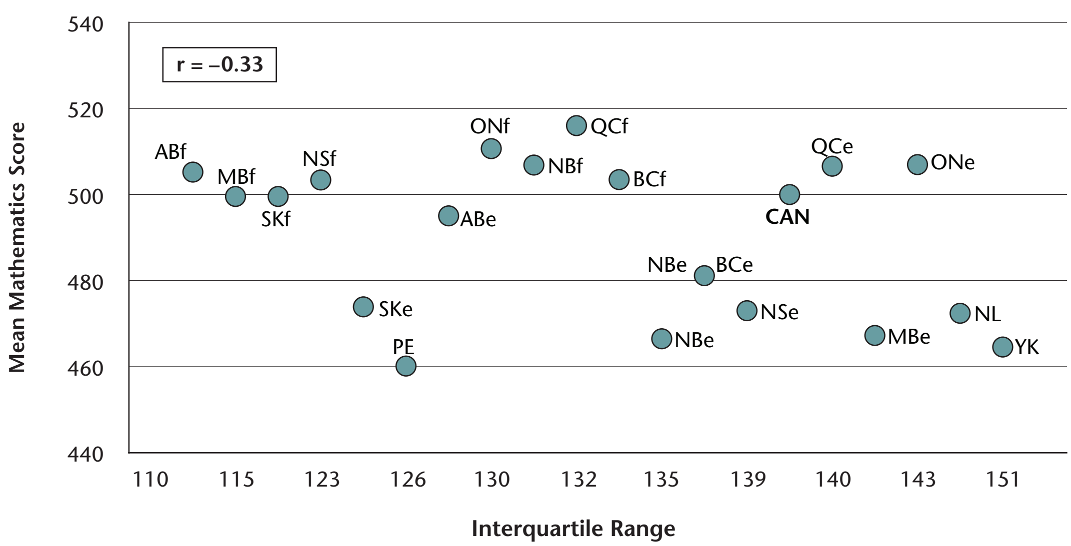


Equity and achievement

An important question to ask is whether there is any relationship between achievement and equity. In particular, it is important to ask whether there is a trade-off between policies designed to promote equity and those designed to promote high achievement. For example, policies designed to encourage differentiation among schools might be seen as leading to higher average achievement but less equity in the system as a whole.

Chart 10-3 is a plot of interquartile range versus mean mathematics score for each population. Overall, the correlation between achievement and interquartile range is negative (r = −0.33). Because the interquartile range is a reverse index of equity, this implies that higher scores are associated with greater equity. A cluster of smaller francophone jurisdictions can be seen with low interquartile ranges (high equity) and relatively high achievement, implying that these populations are closer than others to the desired goal of high achievement and high equity.[22] Four other francophone populations, Ontario French, Quebec French, New Brunswick French, and British Columbia French have above average achievement and medium equity. Quebec English and Ontario English have above average scores and high interquartile ranges (low equity). Some low-performing populations, specifically Manitoba English, Newfoundland and Labrador, and Yukon, combine low achievement with high interquartile ranges (low equity).

Chart 10-3 **Mean mathematics achievement and interquartile achievement range by population**



[22] It would be inappropriate to conclude that the results for these small populations are a consequence of achievement or equity policies in these populations because results from small populations are prone to instability over time. Thus, while reasonably accurate in terms of the standard errors that can be computed from the available data, these results may not be replicated in other assessments. This caution applies generally to all of the PCAP small populations.

Statistical Note

Population Models

The models used in this chapter are a slight variation on the multilevel models used in earlier chapters and explained in the statistical note on multilevel modelling in Chapter 3 (page 48). In this case a separate variable is created for each population, with each student coded as 1 or 0 (referred to as dummy coding) representing whether the student is or is not a member of a specific population. For example, the variable labelled "ONe" contains a 1 if the student is in the Ontario English population or 0 if not.

The counterpart of the simple regression model in earlier analyses is called the "population model." This model contains all but one of the population variables. One population, referred to as the "reference population" is omitted to avoid what is known as a "linear dependency," which prevents the model from being computed. Ontario English has been chosen as the reference population in this case. In order to obtain a coefficient for Ontario English, a second model was run using Quebec French as the reference population. A coefficient in the population model may be interpreted as the mathematics score difference between a specific population and the reference population and is thus referred to as a "population coefficient."

Using the population model as a starting point, the shift in population coefficients may be examined as other variables are added to the model. Comparing these at difference stages shows any differential effects for the populations of the variables entered. For example, if mother's education shows a different effect in, say, Quebec French relative to New Brunswick French, this will appear as a relative difference in how the coefficients for these populations change when mother's education is added to the model.

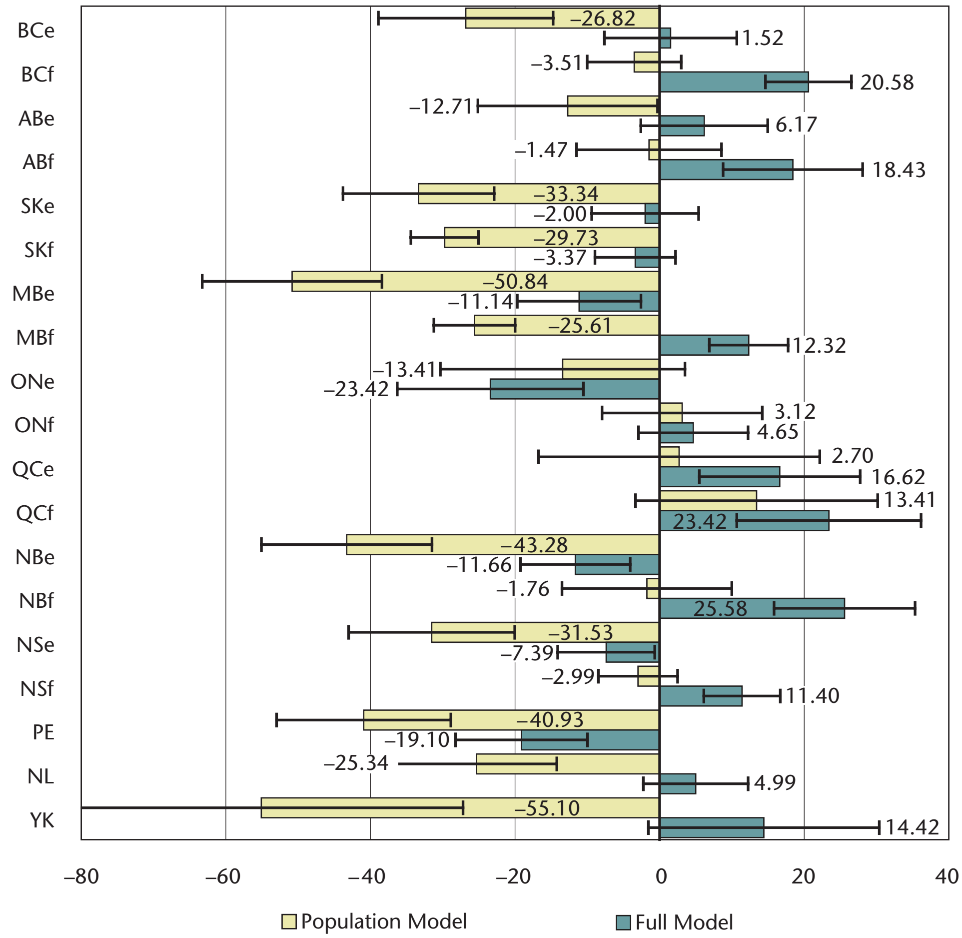
Various intermediate models were computed, corresponding to the variable clusters found in each of Chapters 3 through 9. These were entered cumulatively, with each cluster adding to the previous ones. To distinguish from previous models, the final model is referred to here as a "full model" — one where all variables of interest are controlled. The main interest here is the change in coefficient size for each population in the full model compared to the population model. Overall changes in the predictive power of the model are also examined as each cluster of variables is entered.

The previous chapters presented models for the effects of several clusters of variables on mathematics achievement for Canada as a whole. The concern here is not directly with these effects but rather with how these may act to influence differences in achievement levels across populations. The clusters of variables are:

- Student demographics
- School demographics
- Student attitudes and attributions
- Out-of-school activities
- Student-learning strategies
- Early learning strategies
- Instructional climate
- Special needs
- Challenges to teaching
- Time
- Teaching strategies
- Assessment

Rather than presenting a complex series of "intermediate" models corresponding to the variable clusters used in previous chapters, Chart 10-4 shows the initial "population" model, in which the coefficients represent the effect for each population relative to the reference population (Ontario English for all other populations and Quebec French for Ontario English). This is compared to a "full model" in which the population coefficients represent the effect of each population after controlling for all other variables in the model. The latter variables were selected on the basis of their statistically significant effects in the models presented in previous chapters.

Chart 10-4 **Population coefficients for initial (population) and final (full) models by jurisdiction and language**



In this chart, the coefficients for the population model may be interpreted as the difference in mathematics score between Ontario English and each of the other populations. For Ontario English itself, the population model coefficient is the difference between that population and Quebec French.

The striking feature of Chart 10-4 is the large shift in difference between each population and the reference population when other variables are controlled. In the population model, almost all populations have lower scores relative to Ontario English, with Ontario English having a negative score relative to Quebec French.[23] In the full model, with all other variables controlled, most of the differences between Ontario English and other provinces have been reduced. This implies that these other variables are contributing factors in Ontario's higher performance.

To help make sense of this model, it is useful to look at how the mathematics scores change when variables, such as those appearing to contributing to Ontario's higher performance, are controlled. With all other variables controlled, most of the francophone populations have average mathematics scores significantly higher than those for Ontario English. Only Manitoba English, New Brunswick English, Nova Scotia English, and Prince Edward Island continue to have significantly lower mathematics scores than

[23] The population model coefficients are essentially the same as the comparative results reported in Chapter 2. However, the multilevel model generally gives somewhat wider confidence intervals because of the way in which the error terms are computed.

Ontario English. Even for these populations, the shift itself is statistically significant in a positive direction, implying a significant improvement in their position relative to Ontario English when all other variables are controlled. To take but one example, New Brunswick English has a mathematics score about 43 points lower than Ontario English in the population model. This changes significantly, to about 12 points lower, when other variables are controlled. In the same way, Quebec French moves from being not significantly different in the population model to significantly higher than Ontario English in the full model.[24]

These fairly dramatic shifts in the comparative position of populations raises the obvious question: "Which variables contribute most to the shift?" This question cannot be answered in any simple way. However, a review of the intermediate stages[25] leading to the full model can shed some light on the question. At each of these stages, a cluster of variables was entered, corresponding approximately to the sequence presented in earlier chapters. The difference is that each successive stage was cumulative, adding a new cluster of variables to the previous model.

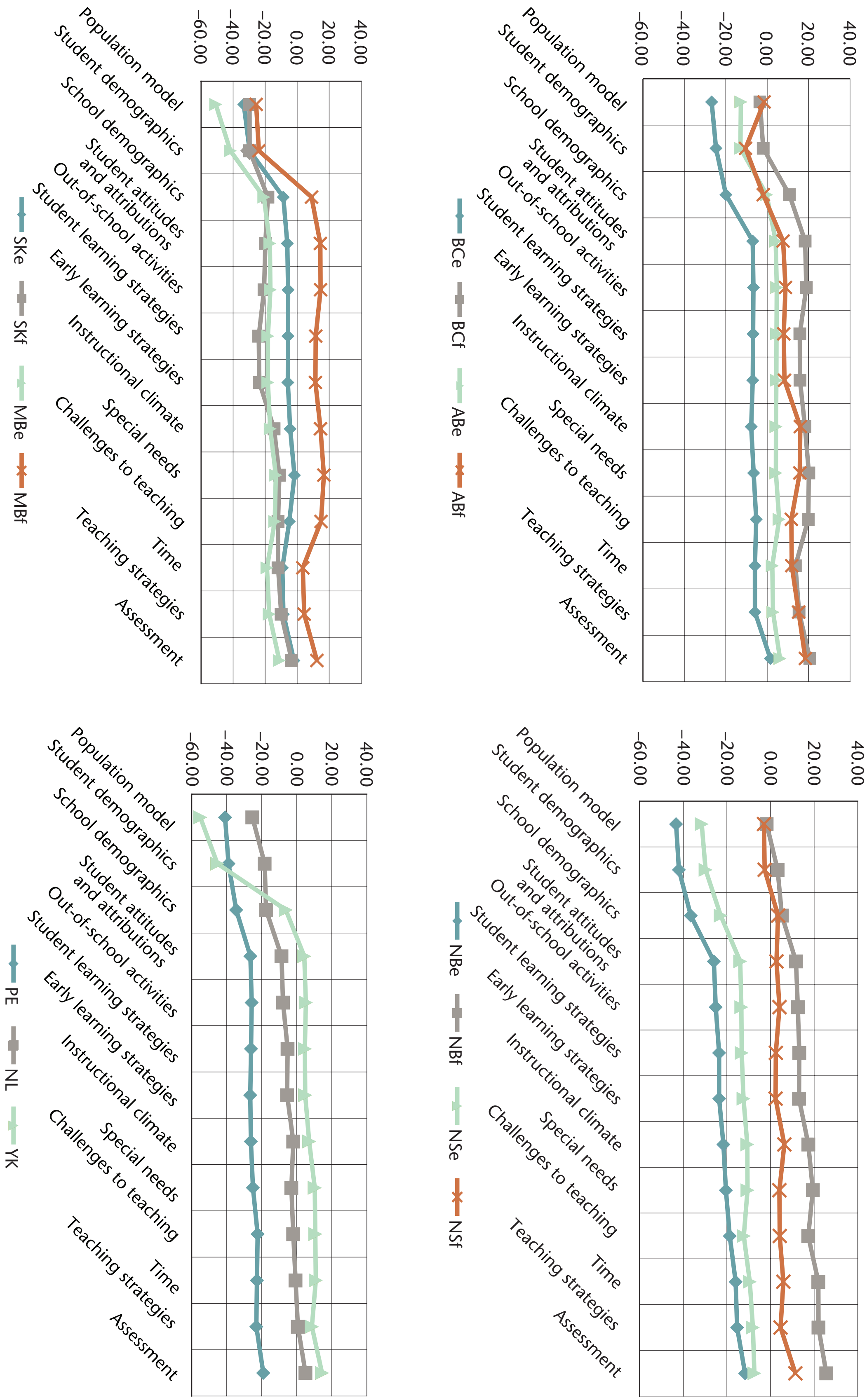
The general pattern in stepwise models of this type is for coefficients to change fairly rapidly in the first few stages and then to level off to the extent that new variables are correlated with those already in the model. Discontinuities in this trend can serve to highlight groups of variables that contribute more than expected to the pattern. This type of reasoning is captured in Chart 10-5. This chart shows the changes at successive stages of the model for all populations whose coefficients changed significantly from the initial (population) to the final (full) model. That last stage is labelled "assessment" in these graphs to identify assessment variables as the ones entered at the last stage.

Keeping in mind that all of the populations shown in the chart had lower scores than Ontario English in the initial population model, the results show that controlling for other variables usually decreases the difference between these populations and Ontario English (brings the difference closer to zero). It can thus be argued that the performance of Ontario English is attributable to differences on most factors that contribute to higher achievement. The most obvious examples are student demographics, school demographics, and student attitudes and attributions, where including these variables in the model substantially diminishes the difference in mathematics performance between Ontario English and other populations. Closer examination at the school level of practises that influence students' perception of and performance in mathematics could be enlightening. Other examples of these variables are being born outside of Canada, the student's highest expected education level, larger schools, and school location in larger communities. All of these show Ontario English at the high end of the scale and a positive association with mathematics achievement. Once these variables are controlled (essentially equalized across populations), the difference in mathematics achievement across jurisdictions diminishes. This provokes some interesting questions for further research.

[24] It is noted that the population model accounts for only about 3 per cent of the total variation in scores, whereas the full model accounts for close to 50 per cent. This implies that differences between populations are a relatively minor source of overall score differences. Chapter 11 gives more detail on proportions of variation accounted for by various clusters of variables.

[25] Coefficients for all of the variables at all intermediate stages of the model are given in Appendix B.

CHART 10-5 **Regression coefficient changes as variable clusters added at each model stage, by population**



The same is true for the student attitude variables. Ontario English tends to be high on the scale for attitude variables that contribute positively to achievement (e.g., confidence in mathematics) and low on those that are negatively associated with achievement (e.g., negative attitudes to mathematics). Again, controlling for the attitude cluster of variables tends to reduce the difference between Ontario and other populations.

Nevertheless, the effects of entering these early stage variables (particularly student demographics, school demographics, and attitudes) differ by population. For example, these variables have a much smaller effect on the population coefficients in the Atlantic provinces than in other regions. For Alberta French, adding student demographics to the model actually increases the difference between that population and Ontario English (makes it more negative). It is not obvious from the descriptive data why this might be so. Another example of this is the time cluster. Adding this cluster to the model increases the score for British Columbia French, Saskatchewan English, Manitoba English, and Manitoba French, relative to Ontario English. One specific variable in this cluster is the average daily percentage of students absent from school. All of these populations have absence rates lower than Ontario English. Once absence rates are equalized, therefore, the expectation is that scores for these populations would fall relative to Ontario, which is what the graph reveals. That is to say, if these populations had absence rates comparable to Ontario English, their mathematics scores would be even lower relative to Ontario English, after controlling for all earlier variable clusters.

The final cluster that might be of interest is the assessment cluster. This cluster includes variables related to assessment processes, test item types, knowledge of and use of rubrics, and others, as described in Chapter 9. For most populations, adding this cluster as the last one making up the full model yields a discontinuity in the direction of increasing the population score relative to Ontario English. As an example of a variable in this cluster, the percentage of Ontario English students who know what a rubric is tends to be relatively high compared to other populations. Knowing what a rubric is positively related to mathematics achievement in the simple regression model (from Chapter 9), but this effect becomes negative in the full model. The result is that knowing what a rubric is results in a score increase for other populations relative to Ontario English.

It is perhaps obvious that these interpretations are complex and that the above comments do no more than scratch the surface of the question of how to account for population differences in mathematics achievement. Compared to the results for 2007 PCAP reading for example, which yielded few significant factors, the results for mathematics suggest that how mathematics is taught and learned in different jurisdictions, and between language groups within jurisdictions, can have a significant impact on mathematics outcomes. Although this point cannot be pursued in more detail here, further investigation of the coefficients presented in the appendices, together with a more detailed analysis of mathematics teaching and learning policies and practices, as revealed by the comparative questionnaire results, would be a useful direction for further research using the PCAP 2010 data base.

11 SUMMARY MODEL, ROBUST EFFECTS, AND FURTHER RESEARCH

The results presented in previous chapters show that a large number of variables are associated with mathematics outcomes, in both positive and negative ways. However, it is also clear from the models that many of these variables are themselves intercorrelated in complex ways, resulting in a situation where many of the simple regression effects are attenuated when other variables are controlled.

The approach to the multivariable analysis taken in earlier chapters was a two-stage one. First, a "simple regression" model was computed for each variable. This gives the raw or "absolute" relationship between that variable and mathematics achievement, without reference to any other variables. At the second "multiple regression" stage, variables within a specific cluster, such as student demographics or attitudes, were controlled. Variables within a cluster tend to be correlated with each other. The change in coefficient for a particular variable, before and after controlling other variables in the same cluster, is an indicator of how much of the initial effect is related to what we may call "interference" from other variables in the cluster. In some cases, controlling for the cluster had little impact on the effect of a specific variable, while in others most of the effect was essentially "absorbed" by the other variables.

To illustrate this point, we can see from Chapter 3 (Chart 3-48) that both books in the home and mother's education are positively related to mathematics achievement in the simple regression model. Even though these two variables are highly correlated (those whose mothers have higher levels of education tend to have more books at home), the effects do not change much when each of these, along with other demographic variables, is added to the model. On the other hand, from Chapter 4 (Chart 4-18), the attitudinal variable "mathematics is easy" has a large positive effect on achievement when taken alone. However, that effect is significantly diminished (though it remains positive) when other attitudinal variables are added to the model. Other attitudinal variables, such as confidence in mathematics, thus combine to reduce the absolute effect of mathematics is easy, giving a smaller effect relative to other variables in the model.

This two-stage approach is extended in this chapter to a more comprehensive model in which variables from all of the categories examined earlier can be controlled. The coefficients in such a model can be considered as unique or residual effects for individual variables, once everything else is controlled. The goal here is to identify these effects that might be considered "robust" in the sense that these remain significant even after many other variables are controlled. These effects might be considered to be of direct policy interest because of their robustness. At the same time, those variables for which significant shifts are found might be considered the ones that warrant further attention to determine the sources of the change. From a statistical modelling perspective, this would require a staged or stepwise approach, in which variables of direct interest are subjected to various stages of control.

Statistical Note

Missing Data and Imputation

Missing data is a significant issue in multiple regression models with a large number of variables. This is because, if data are missing for any one variable for a particular case, that case is deleted before the model is computed. Using a large number of variables can result in a large number of cases being deleted. This can have a significant impact on the results.

The technical solution to the missing data problem is a process known as "imputation." Values are estimated (imputed) for missing data based on patterns found in the non-missing data and the relationships among variables. A simple example of imputation is replacement of all missing values by the mean of the variable. This approach is manageable with small amounts of missing data. However, with larger numbers of missing values, the distribution of values becomes distorted by having a large number of values equal to the mean. In this report, an imputation technique based on multiple regression analysis used to "predict" values for cases with missing values was applied to the data file before the models were computed. Only a small number of cases, in which values were missing for almost all variables, were dropped from the data set when computing the models.

Summary model

The final or summary model used here is an extension of those described in earlier chapters. Instead of entering each cluster separately and looking at effects within each cluster, the clusters were entered cumulatively. At each stage, the predictive power of the model, measured by the proportion of student and school variance accounted for by the model, was extracted along with the coefficients for each variable in the model.

The main interest here is in the final or summary multiple regression model, which includes all of the selected variables. However, examining the predictive power at intermediate stages can give a picture of the relative magnitude of the effects for successive clusters. Also, the change in coefficients across successive stages can help identify how entering new variables influences the effects for variables in the model at earlier stages.

To reduce the number of variables in the model to a manageable level, clusters of variables were selected for inclusion in the model only if they accounted for more than 2 per cent of either the total student-level or total school-level variance. Individual variables were selected only if they were statistically significant in both the original simple and multiple regression models. This reduced the number of variables in the final model from 115 to 57.

Proportions of variance

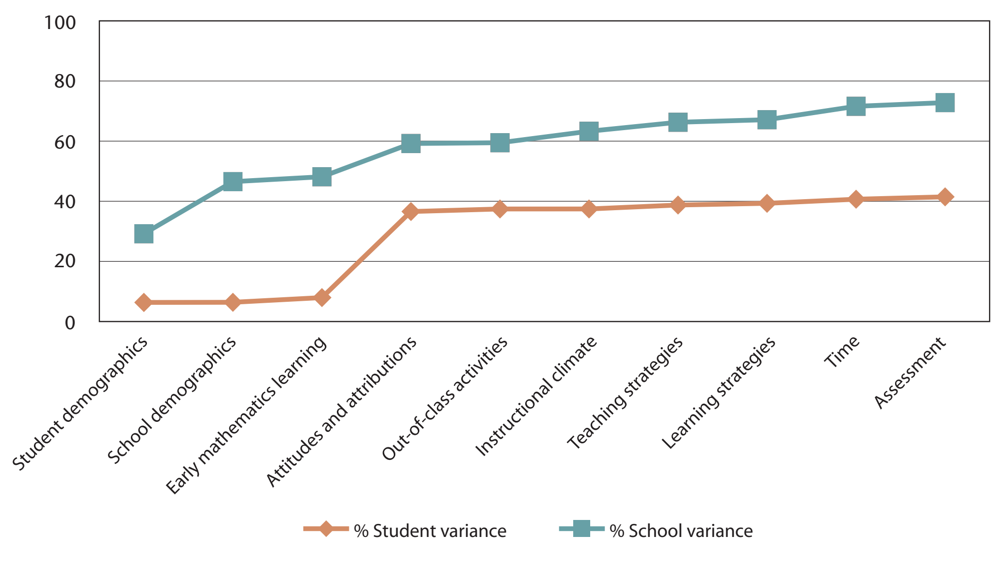
Chart 11-1 shows the proportions of the total student and school variance accounted for by each cumulative stage of the model. These proportions may be interpreted as indicating the predictive power of each stage of the model, at each of the student and school levels. This chart follows a typical pattern of fairly rapid increase in predictive power with the first few variable clusters and a levelling off once more clusters are added.

In this type of "stepwise" approach, the specific proportions of variance depend on the order of entry of the clusters. In this case, the logic of the order of entry was roughly temporal, that is, variable clusters judged to occur first were entered first. As an example, student demographic characteristics were assumed to predate the formation of attitudes and behaviours. These, in turn, predate school exposure and school learning. The assessment cluster was entered last as it was taken to be the culminating event in teaching and learning.

The total variance in the mathematics scores for Canada is simply the square of the standard deviation of the scores, or 10,000. Computing the model with no predictor variables (the null model) shows that approximately 78 per cent of this variance is accounted for by differences between students and the remaining 22 per cent by differences between schools. (The latter includes differences between teachers.) As each cluster is entered, the proportion of the total student and school variance explained by the model is computed.

Chart 11-1 shows these proportions. This shows the increase in predictive power of the model for each successive cluster.

Chart 11-1 **Percentage of student and school variance for model stages**



The first cluster, student demographics, accounts for approximately 6 per cent of the student variance and 29 per cent of the school variance.[26] Student demographic characteristics can thus be said to influence differences between schools as well as differences between students in mathematics achievement. Entering the second cluster, school demographics has a relatively large additional effect on school variance, increasing from 29 per cent to 47 per cent, but almost no effect on student variance. This is to be expected because school demographics cannot be expected to influence student characteristics (though the opposite is logically possible). Entering the early mathematics learning cluster has only a small effect on both variance components. This is likely because the absolute effects of early mathematics learning as shown in Chapter 5 (Charts 5-15 and 5-16) are absorbed or attenuated by the previous entry of student demographic variables.

The greatest single effect on student-level variance is found for the attitude and attribution cluster. This increases the proportion of student variance from 8 per cent to 37 per cent. This cluster also yields about a 10 per cent increase in school-level variance, suggesting that attitudes may be clustered within schools and can influence overall school results as well as individual student results. Beyond the attitude and attribution cluster, only small further increases in predictive power are found. This suggests that the earlier variables may have absorbed much of the effect of the later variables, through their inter-correlations. Nevertheless, an examination of the proportions of variance accounted for by these later clusters, when entered independently, indicate that their predictive power, as separate clusters, is substantially lower than that of the earlier clusters. The small increases in predictive power for the later clusters are thus likely a consequence of the combination of low separate power and the cumulative effect of the early clusters.

The general conclusion from these results is that student and school characteristics, including attitudes, have a larger effect on mathematics achievement than the variables associated with teaching and learning in the school. This might be interpreted as showing that what is done in schools has less influence on learning than the characteristics students bring to schools and the structural features of schools themselves. However, the result also highlights a design issue with cross-sectional surveys of this nature. In effect, the student and school characteristics measured are relatively stable and permanent, whereas the teaching and learning variables tend to be more transient. Most of these variables capture, at best, what might have been done during the school year in which the assessment was conducted and are not necessarily indicative of the student's overall exposure to teaching and learning.

What is of greatest interest in studies of this nature is whether teaching and learning activities can mitigate some of the adverse effects of certain student characteristics. The equity issue examined in the previous chapter illustrates this point. Ideally, schools should attempt to overcome disadvantages or barriers inherent in some student characteristics, thus helping to equalize achievement. The remainder of this chapter can shed some light on this by examining the changes in coefficients in the summary multiple regression model compared to the simple regression model. However, the sources of these changes need to be examined more closely than is possible here before a fuller picture of teaching and learning effects can emerge.

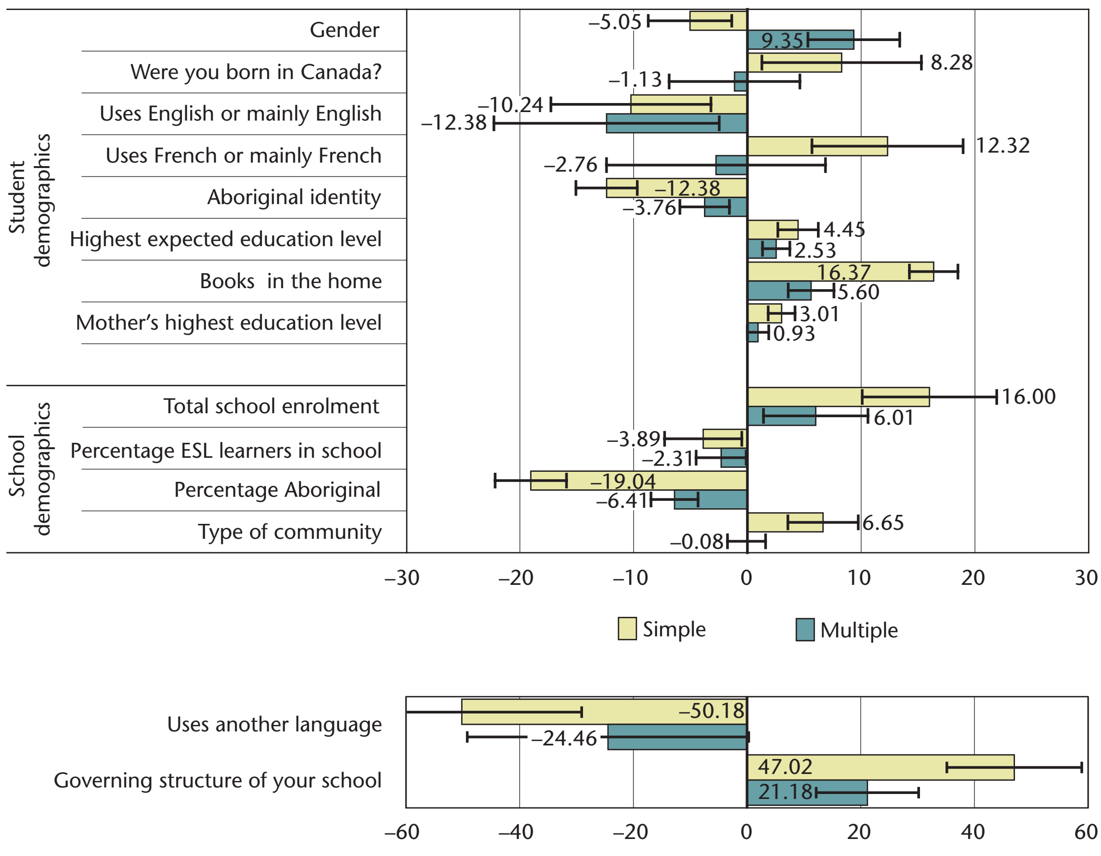
[26] The percentages given in the chart are "percents of percents." For example, the first cluster accounts for 6 per cent of the 78 per cent student variance and 29 per cent of the 22 per cent school variance. It should also be noted that these percentages depend on the order of entry of the clusters. This does not affect the final summary model but does influence the results for intermediate model stages.

The remainder of this chapter looks in more detail at the variable clusters, comparing the simple regression effects to the multiple regression effects in the full summary model. Where large changes are observed, the intermediate model coefficients (Appendix B) are examined briefly to see if a source of change can be seen. However, as already noted, a much more detailed analysis of specific effects is required to more fully explain many of the changes.

Demographic effects

Chart 11-2 shows the simple and multiple regression effects for student and school demographic variables. (Teacher demographics were dropped because they did not meet the entry criteria). Several notable changes from the simple to the multiple regression model are apparent from this chart.

CHART 11-2 **Regression coefficients for student and school demographics [This chart is presented in two sections for clarity of scaling.]**



First, the simple regression gender effect shows that girls perform at a slightly lower level than boys (girls were coded 1 and boys 0 in this analysis). This effect is reversed in the multiple regression model. The intermediate coefficients show that the main contributor to the change is the attitude cluster. Girls generally have less positive attitudes to mathematics than boys. After controlling for attitude, girls are found to have higher mathematics achievement.

This, of course, illustrates the utility of the multiple regression model. Initial simple regression effects may be misleading or may have an obvious explanation once other variables are examined. Nevertheless, caution is needed in these interpretations because the model does not reveal exactly which attitude variables might be important or even whether attitudes are a cause or a consequence of achievement. This level of detail cannot be pursued in an omnibus report such as this one.

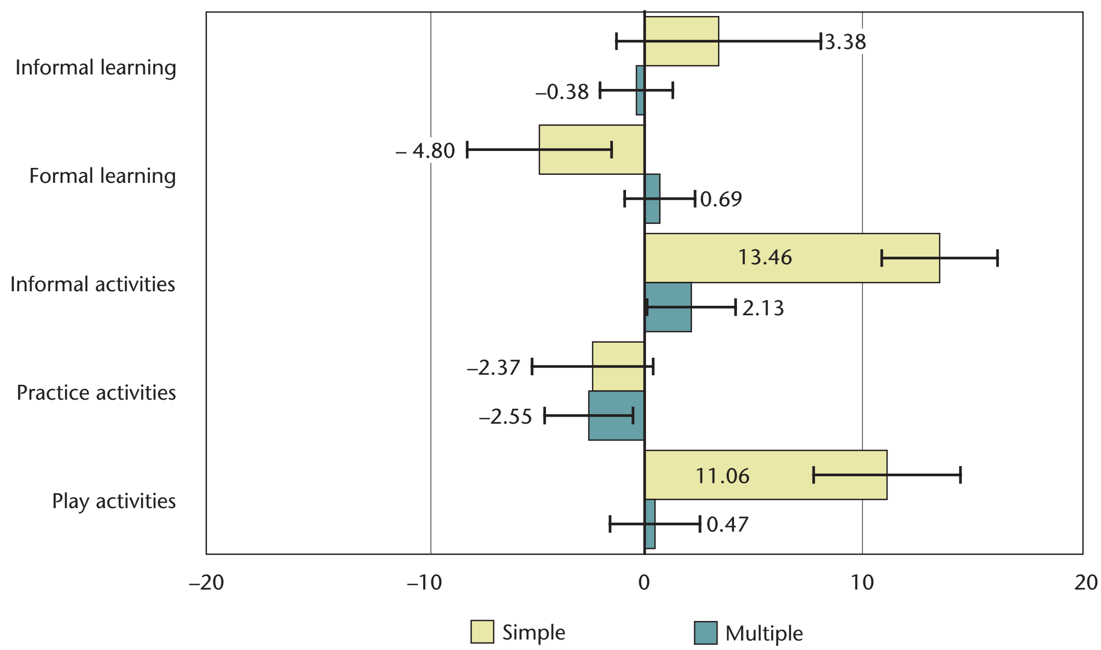
Most of the other effects are attenuated, but not reversed, in the multiple regression model. The change is statistically significant for several variables. The effect on the model of the variable related to Aboriginal identity, which becomes significantly attenuated can be traced to the attitude cluster. Aboriginal students report more negative attitudes to mathematics, and controlling for attitudes improves their performance relative to non-Aboriginal students. Both books in the home and mother's education are also significantly attenuated in the multiple regression model, also with attitudes to mathematics as the main contributor to the change.

For the school variables, there is a significant change in the coefficient for percentage of Aboriginal students in the school. This can be traced in the intermediate models mainly to school demographic variables, suggesting that the type of school in which Aboriginal students are found can contribute to improved achievement. The same is true for type of community. In effect, other school characteristics can contribute to relative improvement in scores for schools in smaller communities. Finally, the large private school advantage found in the simple regression model is highly attenuated in the multiple regression model. Again, the intermediate coefficients show that the main contributor to the change is other school demographic variables.

Early mathematics learning

Chart 11-3 gives the coefficients for early mathematics learning variables. A significant difference can be seen for formal learning, which goes from non-significant to significantly negative when other variables are controlled. This is an example of what is called a "suppressed effect," which cannot be detected in the simple regression model but which emerges when other variables are controlled. Informal learning is also significantly attenuated, but remains significantly positive in both models. Practice activities changes from non-significant to significantly negative. In this case the change in the coefficient is quite small, but the confidence interval is also smaller. Finally, the effect for play activities shifts from significantly positive in the simple regression to non-significant in the multiple regression model. In all of these cases, controlling for attitudes is the main source of the change.

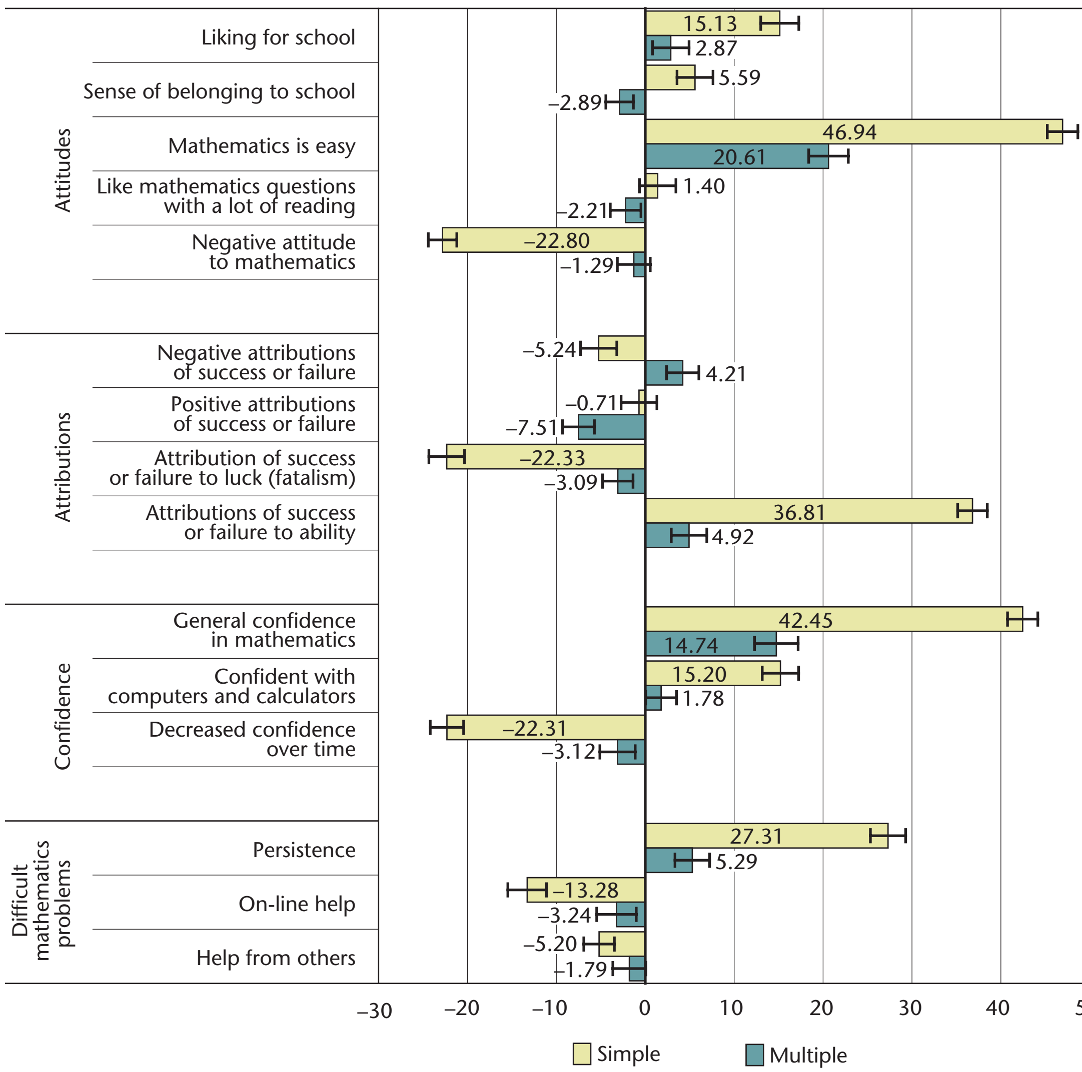
CHART 11-3 **Regression coefficients for early mathematics learning**



Attitudes and attributions

Chart 11-4 gives the coefficients for the attitudes and attributions cluster. The immediately obvious feature of these results is that almost all of the effects are significantly smaller in the multiple than in the simple regression model. However, most also remain statistically significant and in the same direction. This indicates both the pervasiveness of attitudes as determinants of mathematics achievement and the fact that attitudes and attributions can be influenced by other variables. As it happens, however, most of the change can be attributed to the inter-correlations among the attitude variables themselves. The simple regression effects may thus be thought of as multiple measures of the same general trait. Adding further clusters beyond the attitude cluster changes these coefficients only slightly.

CHART 11-4 **Regression coefficients for attitudes and attributions**



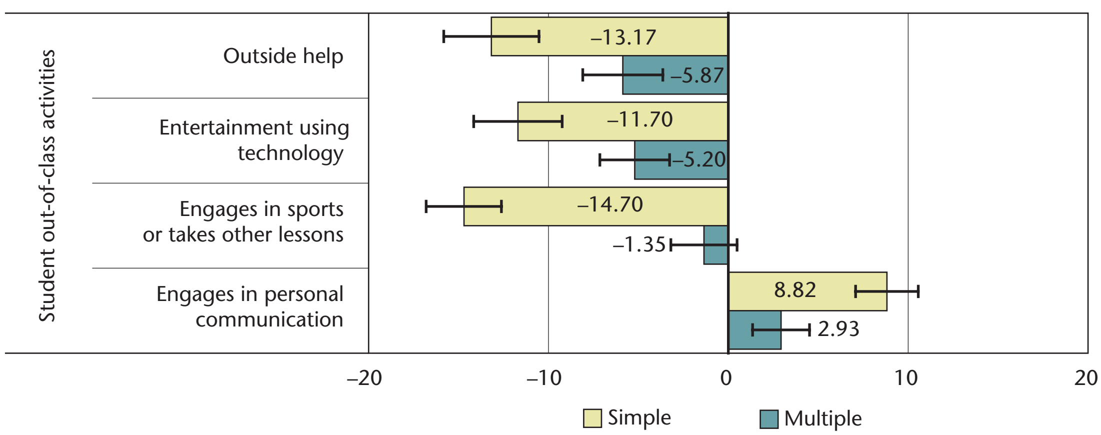
Student out-of-class activities

The next cluster entered into the model was student out-of-class activities. This variable is derived from questions on using the telephone or texting, using a computer for personal reasons, and playing mathematics computer games. Engaging in most of these activities is negatively associated with achievement in the simple regression model. The exception is personal communication. All of these effects are significantly attenuated in the multiple regression model, although all but one (sports or other lessons) remain statistically significant in the same direction in both models.

For this and successive clusters, it is not always possible to clearly identify the source of the change because the effects of clusters previously entered cannot be separated because of the entry sequence. What we can say is that most of the change is related to the previously entered clusters and not to those entered later. As was the case for previous clusters, the most likely source of the shift is attitudes.

The main point to be made here is that the effects of seeking outside help in mathematics (e.g., tutors or extra help at school) and engaging in entertainment using technology (e.g., playing computer games, watching television) remain negative even when all other variables are controlled. The negative effects of such activities are significantly reduced, but not completely offset, by other factors in the model. While the result for seeking outside help may seem counterintuitive, it is possible that those seeking the most outside help are low-performing students. With the data available, it is not possible to examine the impact of such help on individual students. However, this is clearly insufficient to transform low performance to high performance for students seeking such help.

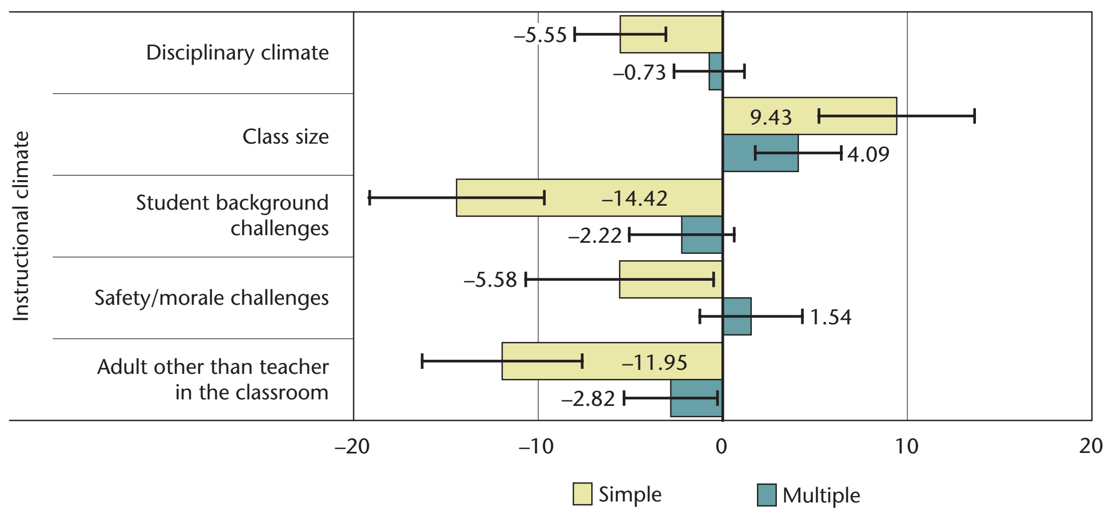
Chart 11-5 **Regression coefficients for student out-of-class activities**



Instructional climate

Chart 11-6 shows the coefficients for instructional climate variables. For three of these variables, disciplinary climate, student background challenges, and safety/morale challenges, the change is from significantly negative to non-significant. The class size effect does not change significantly. This indicates that controlling for a number of variables that are correlated with class size, specifically school size and community size, does not alter the initial finding that students in larger classes perform better in mathematics. The limitation of the model for class size analysis is that it does not control for other variables that might be linked to class size, especially the possibility that lower-performing students might be placed in smaller classes. Nevertheless, even if this is true, the results do little to support the main reason for doing so, namely that this will improve the performance of such students. Finally, having an adult other than the teacher in the classroom remains significantly negative in the multiple regression model. This phenomenon is generally associated with having one or more special-needs students in the classroom, which may complicate classroom life to the detriment of overall achievement.

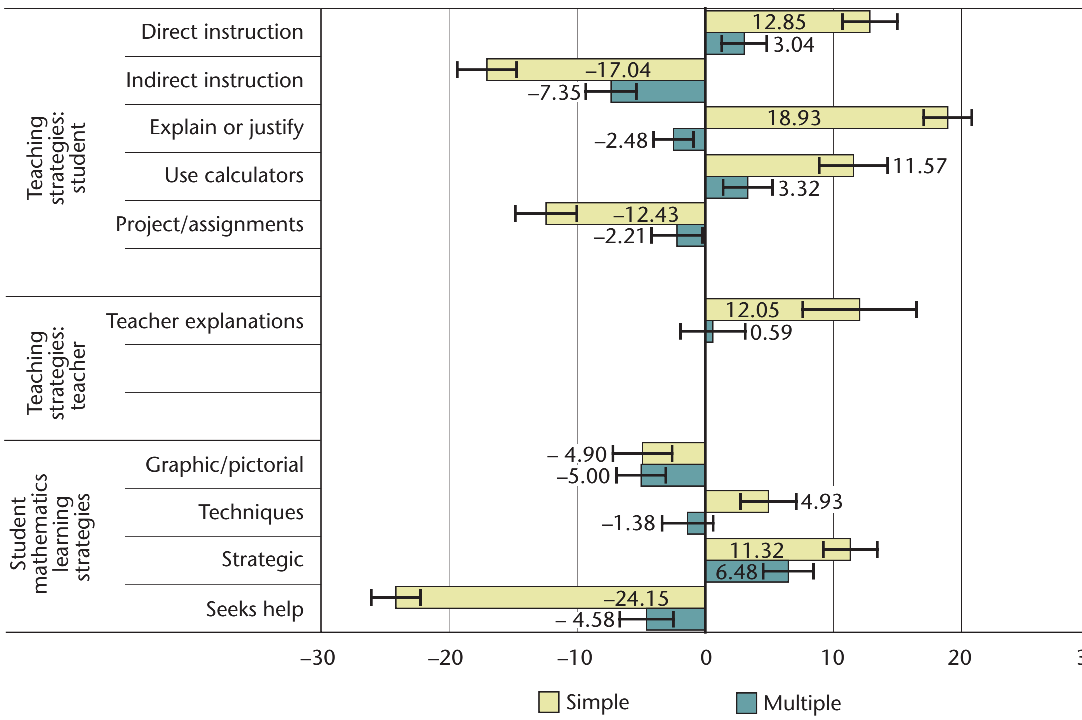
Chart 11-6 **Regression coefficients for instructional climate**



Teaching and learning strategies

Chart 11-7 gives the regression coefficients for teaching and learning strategies. Most of these show significant shifts from the simple to the multiple regression model. However, the effects mostly remain statistically significant in the same direction in both models. Only explaining/justifying as a strategy shows a reversal, from significantly positive to significantly negative. The related teacher variable, teacher explanations, changes from significantly positive to non-significant. The same is true for the strategy of learning mathematics techniques. The use of graphic/pictorial learning strategies is the only variable showing no significant change, showing a relatively small but significant negative effect in both models.

CHART 11-7 **Regression coefficients for teaching and learning strategies**

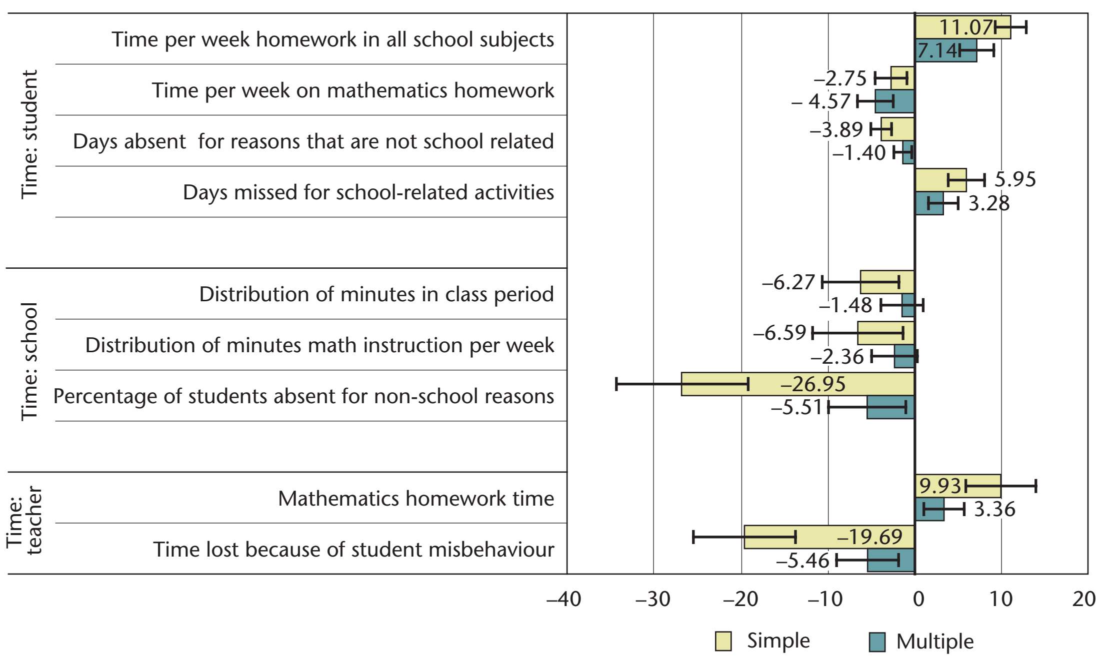


Time allocation and use

Chart 11-8 gives the coefficients for time allocation and use. Most of these do not change significantly from the simple to the multiple regression models. Total homework time per week is positive for both models, while time on mathematics homework shows a positive effect at the teacher level (assigned homework) and a negative effect at the student level (homework actually done). Although the latter is counterintuitive, it is possible that low performing students actually do more mathematics homework in the hope of improving their performance. The data cannot directly answer the question of whether doing more homework actually improves performance for individuals.

Missing school for non–school-related reasons has a negative effect at both the student and the school levels. However, the school-level variable shows a significant decrease in the value when other variables are controlled. This suggests that other school or student characteristics or activities can mitigate the negative effects of average absence rate for the school. Days missed for school-related activities shows a positive effect in both models, suggesting either that higher performing students are the ones engaged in these activities or, alternatively, that these activities actually contribute to mathematics learning.

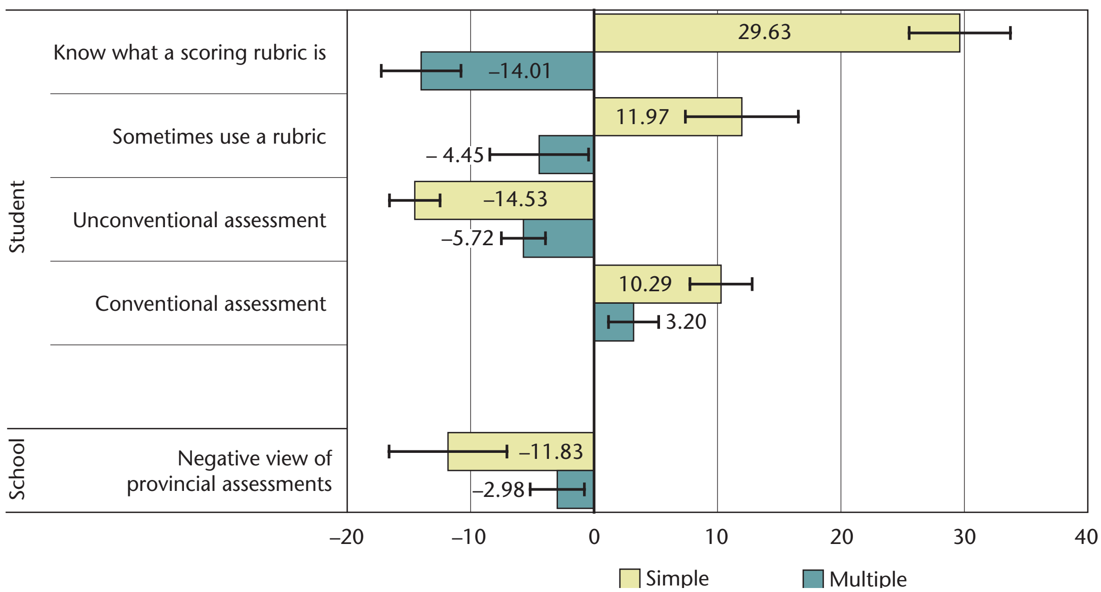
Chart 11-8 **Regression coefficients for time allocation and use**



Assessment

The assessment results in Chart 11-9 show that the effects for knowledge and use of rubrics change from significantly positive to significantly negative from the simple to the multiple regression model. Other variables related to these thus offset the simple regression positive effects. All of the other effects remain significant and in the same direction in both models, with conventional assessment (e.g., tests, homework) showing positive effects and non-conventional assessment (e.g., portfolios, peer-assessment) negative effects. Finally, to the extent that school principals show negative attitudes toward provincial/territorial assessments, their students do less well in mathematics. This effect is significantly reduced in the multiple regression model, although it remains significantly negative.

CHART 11-9 **Regression coefficients for assessment**



Robust effects

Of the many variables that show a relationship to achievement when taken alone, not all remain statistically significant when a large number of other variables are controlled. For purposes of this discussion, those that do may be considered robust enough to have direct implications for policy and practice. These robust effects are identified in Table 11-1.

TABLE 11-1 **Variables that show robust associations with achievement**

Variables associated with higher achievement in mathematics	Variables associated with lower achievement in mathematics
Student level variables	
• Higher expected education level • More books in the home • Higher mother's education • Use of informal early mathematics learning activities • Likes school • Attitude that mathematics is easy • Attribution of success and failure to ability • General confidence in mathematics • Confidence with computers and calculators • Persistence in dealing with difficult mathematics problems • Personal communications • Direct instruction • Uses calculators in mathematics • Strategic approach to mathematics learning • Total homework time • Absence for school-related reasons • Conventional assessment	• Uses English or mostly English in a variety of contexts inside and outside school • Uses a language other than English or French in a variety of contexts inside and outside school • Negative attitudes to mathematics • Attribution of success or failure to luck (fatalism) • Decreased confidence in mathematics ability over time • Use of on-line help in mathematics • Seeks outside help in mathematics • Uses technology for entertainment • Indirect instruction • Instruction by projects/assignments • Graphic/pictorial learning strategies • Seeks help with mathematics • Mathematics homework time • Absence for non–school-related reasons • Unconventional assessment
School/teacher level variables	
• Larger school size • Private school • Larger class size • More teacher-assigned homework time	• Adult other than a teacher in the classroom • Percentage of students absent for non–school-related reasons • Time lost because of student misbehaviour • Negative view of jurisdictional assessments

Further research

It is beyond the scope of an omnibus report such as this to investigate all of the possible links among the variables in the model and how these affect mathematics achievement. The design of PCAP provides for a research phase that would follow the release of the public and contextual reports. For PCAP 2007, selected topics were chosen for more detailed research in a series of three reports that are under way at the time of writing. Research work of this kind can extend the findings of this report by examining theoretical connections among variables and by examining other research on specific issues. The following are some areas that might be worth pursuing in such research.

First, it was not possible in this report to look at other research related to particular variables that might be of policy interest. These include structural features of the school system such as class size and time allocations, as well as teaching and learning variables such as homework, teaching strategies, and assessment practices. Even a cursory examination of previous large-scale surveys such as PCAP 2007 and the SAIP and PISA studies reveals that there are many consistent patterns among the variables affecting achievement in mathematics and other subjects. Before drawing any strong policy conclusions from many of the effects seen here, an effort should be made to determine if these results are consistent with what has been found in other similar studies. The ability to replicate results is key to scientific research, and having consistent results can greatly strengthen any policy decisions that might be made from these results.

Taking class size as an example, large-scale surveys tend to give results consistent with that found here, namely that students in larger classes have higher scores than those in smaller ones, even after controlling for other available variables. However, experimental studies have tended to show the opposite result. Because of the high cost of class size reduction initiatives, it would be desirable to examine the reasons for differences in research findings. For example, the usual working hypothesis in class size research is that smaller classes lead to higher performance. However, the opposite may be true in some circumstances; that is, class size is a consequence, and not a cause, of performance. This would be true, for example, if lower performing students are assigned to smaller classes. Similarly, the experimental class size studies have been mainly in the early years. Class size effects may be different in the intermediate or high school grades, which are the targets of most large-scale assessment surveys.

Class size is only one of many examples for which more detailed examination of the PCAP data, especially in relation to the results of other large-scale assessments, can add value to the assessment and shed greater light on policy issues. Issues such as homework time, absence from school, teaching and learning strategies (especially direct versus indirect instruction in mathematics), and assessment practices are of direct policy interest because these are the things that schools and school systems can do something about. These issues all can be investigated further using large-scale assessment data together with research from other sources.

APPENDIX A: POPULATION MODEL COEFFICIENTS

	Population model		Student demographics		School demographics		Student attitudes and attributions		Out-of-school activities		Student learning strategies		Early learning strategies		Instructional climate		Special needs		Challenges to teaching		Time		Teaching strategies		Assessment: Full Model	
Block	Coeff	SE	Coeff	SE	Coeff	SE	Coeff	SE	Coeff	SE	Coeff	SE	Coeff	SE	Coeff	SE	Coeff	SE	Coeff	SE	Coeff	SE	Coeff	SE	Coeff	SE
BCe	–26.82	6.38	–24.71	5.70	–19.92	6.14	–6.98	5.35	–6.64	5.39	–6.80	5.23	–6.97	5.24	–7.72	5.19	–6.54	5.37	–5.30	5.33	–5.92	4.88	–5.91	4.62	1.52	4.84
BCf	–3.51	8.12	–1.91	8.74	10.70	9.42	18.33	7.28	18.89	7.18	15.76	7.26	15.83	7.18	18.24	8.38	20.12	8.11	19.83	7.72	13.63	6.11	15.40	6.24	20.58	7.42
ABe	–12.71	6.63	–12.92	5.88	–0.28	5.57	4.11	4.74	4.63	4.72	4.50	4.53	4.43	4.52	4.25	4.47	4.21	4.44	5.78	4.41	2.48	4.27	2.81	4.24	6.17	4.68
ABf	–1.47	10.23	–10.46	10.27	–1.90	11.14	7.70	9.95	8.83	9.97	8.00	9.40	8.35	9.28	15.99	10.43	15.74	11.13	11.67	11.19	11.86	11.06	15.17	10.68	18.43	9.84
SKe	–33.34	6.04	–29.68	5.23	–8.65	4.98	–6.16	4.55	–5.67	4.56	–5.73	4.40	–5.81	4.39	–4.38	4.52	–1.68	4.66	–4.92	4.47	–9.23	4.20	–8.59	4.05	–2.00	4.27
SKf	–29.73	6.81	–29.68	5.23	–18.28	12.85	–19.87	6.44	–20.65	6.84	–23.99	8.90	–23.63	9.19	–14.42	9.83	–11.14	8.23	–11.93	7.54	–11.90	8.33	–9.95	8.18	–3.37	8.11
MBe	–50.84	7.10	–41.74	6.10	–20.78	5.82	–17.05	5.10	–16.72	5.11	–18.27	4.93	–18.28	4.91	–16.80	4.94	–13.33	5.17	–13.98	5.12	–18.86	5.02	–17.44	4.98	–11.14	4.88
MBf	–25.61	8.11	–23.93	8.04	8.98	10.80	14.48	7.85	14.63	7.92	11.58	7.63	11.41	7.67	14.59	8.26	16.69	8.29	14.91	7.93	3.52	7.29	4.42	7.24	12.32	7.86
ONe	–13.41	8.73	–17.10	9.47	4.54	9.91	–20.34	7.75	–20.09	7.76	–18.28	7.53	–18.25	7.55	–16.84	7.40	–15.89	7.41	–20.90	7.34	–25.11	6.64	–23.60	6.53	–23.42	6.64
ONf	3.12	7.18	4.88	7.19	3.53	6.29	–0.35	5.42	14.63	7.92	11.58	7.63	–0.50	5.27	2.62	5.49	3.86	5.41	1.78	5.34	2.20	5.01	3.54	4.92	4.65	4.91
QCe	2.70	10.86	3.25	10.07	–1.74	9.50	13.08	8.25	13.15	8.24	13.98	7.97	14.16	7.92	15.09	7.92	15.47	7.68	19.19	7.52	20.31	6.77	17.37	6.64	16.62	6.21
QCf	13.41	8.73	17.10	9.47	–4.54	9.91	20.34	7.75	20.09	7.76	18.28	7.53	18.25	7.55	16.84	7.40	15.89	7.41	20.90	7.34	25.11	6.64	23.60	6.53	23.42	6.64
NBe	–43.28	7.17	–41.99	6.22	–36.53	6.03	–25.96	4.99	–25.11	4.99	–23.56	4.88	–23.55	4.86	–21.59	4.92	–20.44	4.84	–18.69	4.73	–15.98	4.59	–15.30	4.56	–11.66	4.61
NBf	–1.76	8.24	3.27	9.48	5.22	8.64	11.68	6.63	12.53	6.69	13.20	6.60	13.08	6.62	17.35	7.06	19.42	7.12	17.32	6.72	21.94	7.11	22.02	6.96	25.58	6.84
NSe	–31.53	6.73	–29.76	5.79	–22.82	5.30	–13.89	4.40	–13.36	4.37	13.20	6.60	–12.38	4.21	–10.64	4.28	–10.54	4.24	–12.24	4.08	–9.50	3.75	–8.07	3.65	–7.39	3.93
NSf	–2.99	9.25	–2.68	9.52	3.48	8.78	2.74	6.64	4.12	6.60	2.44	6.54	2.33	6.41	6.34	7.14	4.07	7.32	4.24	7.62	5.90	8.21	4.62	8.53	11.40	8.94
PE	–40.93	7.28	–38.97	6.24	–34.59	5.79	–26.46	5.02	–25.73	5.03	–26.09	5.24	–26.54	5.22	–26.19	5.85	–25.01	5.60	–22.34	5.74	–22.73	5.29	–23.08	5.08	–19.10	5.51
NL	–25.34	7.26	–18.31	6.46	–17.54	6.02	–8.70	4.91	–7.85	4.86	–5.21	4.72	–5.51	4.70	–1.98	4.97	–3.06	4.98	–2.00	4.84	–0.68	4.76	0.65	4.72	4.99	4.73
YK	–55.10	22.00	–45.33	18.94	–6.28	19.64	4.30	16.54	5.20	16.88	4.61	15.84	4.82	15.71	7.08	16.24	10.28	15.89	10.58	14.43	10.87	11.97	8.64	11.71	14.42	12.59

NOTE: All coefficients except those for ONe represent change in mathematics score relative to ONe as variable clusters are added to the model cummulatively.

APPENDIX B: FULL MODEL COEFFICIENTS

Variable clusters	Factors or variables	Simple regression		Student demographic		School demographic		Student early mathematics learning		Student attitudes and attributions		Student out-of-class activities		Instructional climate		Teaching strategies		Student learning strategies		Time		Assessment	
		Coeff	SE	Coeff	SE	Coeff	SE	Coeff	SE	Coeff	SE	Coeff	SE	Coeff	SE	Coeff	SE	Coeff	SE	Coeff	SE	Coeff	SE
Student demographics	Gender	−5.05	1.87	−7.35	1.84	−7.33	1.83	−7.78	1.87	10.13	1.79	13.67	1.88	13.59	1.89	10.67	1.96	10.73	1.96	10.15	2.07	9.35	2.06
	Were you born in Canada?	8.28	3.58	8.23	3.88	7.71	3.93	6.68	3.94	−1.46	2.83	−0.72	2.77	−0.59	2.74	−0.24	2.65	−0.60	2.69	−1.20	2.82	−1.13	2.93
	Uses English or mainly English	−10.24	3.59	−10.64	7.51	−6.33	7.63	−7.35	8.01	−9.08	6.46	−9.63	6.46	−10.38	6.42	−11.80	6.23	−12.33	6.47	−11.68	5.26	−12.38	5.05
	Uses French or mainly French	12.32	3.39	3.84	6.77	3.80	6.83	2.49	7.09	−3.85	5.85	−4.13	5.79	−4.21	5.78	−5.18	5.59	−6.14	5.82	−3.33	5.08	−2.76	4.91
	Uses another language	−50.18	10.87	−38.51	15.38	−38.40	15.21	−39.92	15.02	−29.35	10.98	−29.44	11.28	−29.80	11.31	−27.81	11.10	−27.24	11.23	−22.16	11.39	−24.46	12.63
	Aboriginal ancestry	−12.38	1.38	−10.56	1.38	−8.75	1.41	−9.21	1.39	−5.10	1.11	−5.09	1.11	−4.95	1.11	−4.12	1.10	−4.16	1.11	−3.87	1.12	−3.76	1.11
	Highest expected education level	4.45	0.91	3.42	0.86	3.41	0.86	3.53	0.85	2.66	0.67	2.56	0.65	2.53	0.65	2.45	0.63	2.27	0.63	2.24	0.63	2.53	0.61
	Books in the home	16.37	1.09	16.02	1.22	16.12	1.21	15.21	1.23	8.04	1.07	7.57	1.04	7.50	1.04	6.96	1.02	6.91	1.03	6.23	1.03	5.60	1.02
	Mother's highest education level	3.01	0.60	1.62	0.62	1.53	0.62	1.48	0.61	0.90	0.51	0.84	0.48	0.83	0.48	0.75	0.47	0.86	0.47	0.84	0.48	0.93	0.48
School demographics	Total school enrolment	16.00	3.02			9.10	2.66	9.10	2.61	9.60	2.41	9.82	2.42	7.37	2.48	6.18	2.34	6.22	2.32	7.75	2.36	6.01	2.34
	Governing structure of your school	47.02	6.06			25.61	5.45	25.76	5.43	27.58	4.98	28.13	4.99	27.15	4.87	23.88	4.91	24.26	4.85	21.38	4.51	6.01	2.34
	Percentage ESL learners in school	−3.89	1.73			−4.86	1.61	−4.72	1.59	−3.63	1.40	−3.63	1.40	−2.95	1.31	−2.59	1.26	−2.64	1.23	−2.39	1.12	−2.31	1.12
	Percentage Aboriginal	−19.04	1.60			−13.31	1.30	−13.13	1.30	−11.55	1.19	−11.55	1.18	−9.65	1.13	−9.17	1.11	−9.27	1.09	−7.01	1.05	−6.41	1.06
	Type of community	6.65	1.58			3.92	1.38	3.83	1.35	2.65	1.10	2.73	1.09	1.90	1.07	1.93	1.03	1.80	1.01	0.08	0.90	−0.08	0.85
Student early mathematics learning	Informal ways of early mathematics learning	3.38	2.38					7.26	1.18	−0.32	0.87	−0.23	0.89	−0.26	0.89	−0.07	0.84	−0.49	0.83	−0.51	0.85	−0.38	0.85
	Formal ways of initial mathematics learning	−4.80	1.68					−6.25	1.00	0.45	0.81	0.43	0.82	0.46	0.82	0.31	0.78	0.37	0.79	0.42	0.80	0.69	0.82
	Informal mathematics activities before starting school	13.46	1.35					6.56	1.10	3.33	0.90	3.21	0.90	3.25	0.91	2.74	0.94	2.20	0.94	2.32	1.00	2.13	1.03
	Practice mathematics activities before starting school	−2.37	1.41					−5.96	0.84	−5.54	0.86	−4.35	0.90	−4.31	0.89	−3.06	0.94	−2.29	1.00	−2.80	1.02	−2.55	1.03
	Play mathematics activities before starting school	11.06	1.71					3.86	1.24	1.27	1.03	0.41	0.98	0.37	0.99	0.76	0.99	0.49	0.99	0.42	1.04	0.47	1.05

Variable clusters	Factors or variables	Simple regression		Student demographic		School demographic		Student early mathematics learning		Student attitudes and attributions		Student out-of-class activities		Instructional climate		Teaching strategies		Student learning strategies		Time		Assessment	
		Coeff	SE	Coeff	SE	Coeff	SE	Coeff	SE	Coeff	SE	Coeff	SE	Coeff	SE	Coeff	SE	Coeff	SE	Coeff	SE	Coeff	SE
Student attitudes and attributions	Liking for school	15.13	1.09							4.63	1.03	4.62	1.02	4.48	1.02	3.72	1.04	3.92	1.04	3.24	1.08	2.87	1.05
	Sense of belonging to school	5.59	1.03							–3.75	0.75	–3.37	0.81	–3.41	0.80	–2.76	0.78	–2.72	0.79	–2.93	0.81	–2.89	0.79
	Mathematics is easy	46.94	0.88							23.52	1.24	23.18	1.26	23.05	1.24	22.18	1.22	21.51	1.22	21.42	1.18	20.61	1.14
	Like mathematics questions with a lot of reading	1.40	1.05							–2.44	0.88	–2.08	0.88	–2.13	0.87	–2.11	0.85	–2.33	0.85	–2.23	0.87	–2.21	0.88
	Negative attitude to mathematics	–22.80	0.82							–2.57	0.92	–2.36	0.92	–2.31	0.92	–1.61	0.92	–1.57	0.87	–1.33	0.91	–1.29	0.94
	General confidence in mathematics	42.45	0.87							16.47	1.18	15.97	1.21	15.96	1.18	15.68	1.21	15.32	1.19	15.00	1.26	14.74	1.25
	Confident with computers and calculators	15.20	1.04							2.83	0.86	2.45	0.81	2.43	0.82	2.01	0.81	2.26	0.80	2.29	0.87	1.78	0.88
	Decreased confidence over time	–22.31	0.96							–2.74	0.99	–2.95	1.02	–2.95	1.04	–2.77	0.98	–3.05	0.96	–2.99	0.96	–3.12	1.01
	Negative attributions of success or failure	–5.24	1.04							4.79	0.88	4.65	0.88	4.70	0.88	4.23	0.87	4.28	0.88	4.25	0.92	4.21	0.93
	Positive attributions of success or failure	–0.71	1.03							–8.90	0.91	–8.04	0.91	–8.07	0.91	–7.64	0.90	–7.27	0.90	–7.80	0.90	–7.51	0.91
	Attribution of success or failure to luck (fatalism)	–22.33	1.02							–4.80	0.84	–4.66	0.83	–4.59	0.83	–3.30	0.85	–3.55	0.86	–3.43	0.86	–3.09	0.87
	Attributions of success or failure to ability	36.81	0.85							8.95	0.99	6.66	0.97	6.66	0.97	6.35	0.98	5.69	1.02	5.09	1.03	4.92	1.02
	Persistence	27.31	1.01							7.82	0.88	8.09	0.89	8.03	0.89	6.58	0.91	5.81	0.98	5.53	1.00	5.29	0.99
	On–line help	–13.28	1.11							–7.04	1.05	–6.20	1.07	–6.17	1.11	–4.39	1.06	–3.28	1.12	–3.23	1.13	–3.24	1.14
	Help from others	–5.20	0.88							–2.12	0.85	–2.15	0.87	–2.18	0.86	–2.54	0.89	–1.82	0.93	–1.94	0.95	–1.79	0.95
Student out-of-class activities	Outside help	–13.17	1.35									–7.67	0.99	–7.64	1.00	–6.59	1.01	–5.37	1.05	–6.09	1.12	–5.87	1.14
	Entertainment using technology	–11.70	1.25									–5.82	0.97	–5.91	0.99	–5.68	0.98	–5.18	0.96	–5.09	0.96	–5.20	0.99
	Engages in sports or takes other lessons	–14.70	1.06									–3.00	0.85	–2.89	0.84	–3.09	0.85	–2.69	0.85	–1.52	0.92	–1.35	0.94
	Engages in personal communication	8.82	0.89									2.49	0.79	2.45	0.78	2.59	0.76	2.42	0.75	2.86	0.79	2.93	0.81

Variable clusters	Factors or variables	Simple regression		Student demographic		School demographic		Student early mathematics learning		Student attitudes and attributions		Student out-of-class activities		Instructional climate		Teaching strategies		Student learning strategies		Time		Assessment	
		Coeff	SE	Coeff	SE	Coeff	SE	Coeff	SE	Coeff	SE	Coeff	SE	Coeff	SE	Coeff	SE	Coeff	SE	Coeff	SE	Coeff	SE
Instructional climate	Disciplinary climate	–5.55	1.26											–1.36	0.94	–1.10	0.88	–1.19	0.83	–1.16	0.93	–0.73	0.97
	Class size	9.43	2.15											5.55	1.39	4.56	1.35	4.66	1.31	4.49	1.21	4.09	1.19
	Student background challenges	–14.42	2.43											–5.11	1.59	–4.68	1.56	–4.64	1.52	–2.47	1.51	–2.22	1.45
	Safety/morale challenges	–5.58	2.60											2.90	1.63	3.26	1.56	2.99	1.52	2.15	1.46	1.54	1.42
	Adult other than teacher for special needs	–11.95	2.21											–4.43	1.42	–4.04	1.33	–3.99	1.32	–2.91	1.35	–2.82	1.29
Teaching strategies	Direct instruction	12.85	1.09													4.20	0.85	4.48	0.90	3.59	0.89	3.04	0.90
	Indirect instruction	–17.04	1.18													–9.17	0.95	–8.40	0.96	3.59	0.89	–7.35	1.01
	Explain/justify	18.93	0.96													–3.74	0.77	–3.47	0.77	–3.02	0.81	–2.48	0.79
	Use calculators	11.57	1.36													3.66	0.95	3.68	0.97	3.39	0.94	3.32	0.98
	Project assignments	–12.43	1.22													–4.06	1.00	–3.60	1.00	–2.78	1.01	–2.21	1.01
	Textbook assignments	12.05	2.27													1.68	1.43	1.72	1.42	1.07	1.34	0.59	1.29
Student learning strategies	Graphic/pictoral mathematics learning	–4.90	1.18															–5.07	0.92	–5.12	0.97	–5.00	0.98
	Mathematics learning techniques	4.93	1.11															–1.41	0.97	–1.52	1.00	–1.38	1.02
	Strategic approach to mathematics learning	11.32	1.07															6.16	0.95	6.48	0.98	6.48	1.00
	Use of Internet/tutor in mathematics learning	–24.15	0.98															–4.72	1.07	–4.86	1.09	–4.58	1.07
Time	Time per week homework in all school subjects	11.07	0.91																	7.27	0.96	7.14	1.00
	Time per week on mathematics homework	–2.75	0.94																	–4.43	1.01	–4.57	1.05
	Days absent for reasons that are not school related	–3.89	0.61																	–1.37	0.51	–1.40	0.53
	Days missed for school–related activities	5.95	1.07																	3.42	0.86	3.28	0.87
	Distribution of minutes in class period	–6.27	2.26																	–2.18	1.29	–1.48	1.25

Variable clusters	Factors or variables	Simple regression		Student demographic		School demographic		Student early mathematics learning		Student attitudes and attributions		Student out-of-class activities		Instructional climate		Teaching strategies		Student learning strategies		Time		Assessment	
		Coeff	SE	Coeff	SE	Coeff	SE	Coeff	SE	Coeff	SE	Coeff	SE	Coeff	SE	Coeff	SE	Coeff	SE	Coeff	SE	Coeff	SE
	Distribution of minutes math instruction per week	–6.59	2.66																	–2.53	1.37	–2.36	1.35
	Percentage of students absent for non–school reasons	–26.95	3.92																	–5.73	2.26	–5.51	2.28
	Mathematics homework time	9.93	2.06																	3.59	1.21	3.36	1.19
	Lose time because of student misbehaviour																			–5.24	1.80	–5.46	1.83
Assessment	Know what a scoring rubric is (reverse)	29.63	2.09																			–14.01	1.65
	Sometimes use a rubric (reverse)	11.97	2.33																			–4.45	2.04
	Unconventional assessment	–14.53	1.05																			–5.72	0.91
	Conventional assessment (tests, homework)	10.29	1.29																			3.20	1.04
	Negative view of provincial assessments	–11.83	2.44																			–2.98	1.12