INTRODUCTION

A CONTEXT FOR THIS REPORT

This document is the report to the public on the results of the pan-Canadian assessment of writing achievement for 13-year-old and 16-year-old students administered in the spring of 2002 by the Council of Ministers of Education, Canada (CMEC) as a part of the ongoing School Achievement Indicators Program (SAIP).

SAIP is a cyclical program of pan-Canadian assessments of student achievement in reading and writing, mathematics, and science that has been conducted by CMEC since 1993.

The SAIP Writing III Assessment (2002) is the third in a series of writing assessments. Other writing assessments were administered in 1994 and 1998, but their results cannot be compared with those from 2002.

In addition to presenting the results for Canada and for the individual jurisdictions, this public report outlines the conceptual framework and criteria upon which the test is based. As well, it describes the specific development and modification of the test instruments. A preliminary discussion of the data is included, as are the results of a national expectations-setting process, in which actual student results are compared to expectations set by a pan-Canadian panel.

A more detailed statistical analysis of data and a more detailed discussion of methodology will be found in a forthcoming technical report to be released by CMEC.

An important aspect of this assessment is the aggregation of contextual data reported by students on the opportunities they have had to learn to write, on their attitudes toward language and learning, and on their interests and activities related to literacy. Additional contextual information was gathered from school principals and language arts teachers. A sampling of this information is included in this report, while more information and a detailed discussion will be found in the accompanying report, *Student Writing: The Canadian Context.*

Box 1

SAIP Reports

Three reports will be released for this assessment.

- This public report, intended to give a summary of results and how they were obtained.
- An addendum, *Student Writing: The Canadian Context*, with detailed analysis of the data from student, teacher, and school questionnaires, released in conjunction with this report.
- A technical report, which usually follows the public report by several months and contains both a more detailed description of development and administration and a more complete and detailed data set. This report is intended for researchers and education officials.
- The data are available for research on request.

Both public reports will be available on the CMEC Web site at **www.cmec.ca**.

THE SCHOOL ACHIEVEMENT INDICATORS PROGRAM

Canadians, like citizens of many other countries, want their children to have the best educational preparation possible. Consequently, they have asked how well our education systems prepare students for lifelong learning and for the global economy. Do our students have the thinking skills, the problem-solving skills, and the communication skills to meet the challenges of their future?

To provide jurisdictions with a wider pan-Canadian and international context in which to answer these significant questions, ministries[1] of education have participated in a variety of studies since the mid-eighties. At the international level, through CMEC, they took part in the International Educational Indicators Program of the Organisation for Economic Co-operation and Development (OECD), including the Programme for International Student Assessment of 2000 (PISA), involving 32 nations. Individual jurisdictions participated in various achievement studies such as the IEA Reading Literacy Study, the Third International Mathematics and Science Study (TIMSS), and the Progress in International Reading Literacy Study (PIRLS). In addition, most jurisdictions conduct their own evaluations of students at different stages in their schooling.

Since all ministers of education strive to bring the highest degree of effectiveness and quality to their systems, they recognize a need for collective action to assess these systems in a Canadian context. To the extent that all Canadian students learn common skills in the key subject areas of language, mathematics, and science, these subjects provide a common ground for performance assessment on a pan-Canadian level. Consequently, achievement in these school subjects can serve as a useful indicator of an education system's performance.

Therefore, in 1989, CMEC initiated the School Achievement Indicators Program (SAIP). In December 1991, in a memorandum of understanding, the ministers agreed to assess the achievement of 13- and 16-year-olds in reading, writing, and mathematics. In September 1993, the ministers further agreed to include the assessment of science. They decided to administer the same assessment instruments to the two age groups to study the change in student knowledge and skills due to the additional years of instruction. The information collected through the SAIP assessments would be used by each jurisdiction to set educational priorities and plan program improvements.

It was decided that the assessments would be administered in the spring of each year according to the schedule provided in Table 1.

Table 1		
SAIP Assessment Schedule		
Mathematics	*Reading and Writing*	*Science*
1993	1994	1996
1997	1998	1999
2001	2002 (Writing only)	2004

Copies of reports for assessments administered since 1996 can be found in both official languages through the CMEC Web site. For earlier reports, contact CMEC directly at **www.cmec.ca**.

FEATURES OF SAIP ASSESSMENTS

A Brief History of the Development of Assessment Materials

All the provinces and territories were involved in developing the original writing assessment materials beginning in December of 1990. Using a cooperative, consensus-driven approach, the development went forward under the guidance of a consortium team from Alberta, Ontario, and Quebec. Ministries of education reviewed assessment proposals, draft criteria, and assessment materials according to their criteria and the proposed assessment framework. All of the suggestions and concerns informed the revisions.

For the writing assessment, English and French descriptive criteria were identified to design the assessment and scoring rubric. These criteria and the rubric grew out of examining writing produced by 13- and 16-year-olds in classrooms. Consultation over a two-year period was then expanded to include ministry personnel, educators at all levels, and interested members of the general public. Concerns and suggestions directed multiple revisions of the criteria and assessment design.

---

[1] In this report, "ministry" means "department" as well, and "jurisdiction" means both "province" and "territory."

In 1992, the writing assessment was field tested across Canada in both official languages. English- and French-speaking teachers from several jurisdictions assessed the field tests, confirming the appropriateness and range of difficulty for the resource materials, task, instructions, administrative procedures, criteria, and scoring procedures. Teachers with students writing the field test reviewed the instructions, administrative procedures, time, criteria, student questionnaires, student resource booklets, and writing tasks. Developers also considered student comments on these aspects of the field test. This information formed the basis for further decisions and revisions, and the assessment was administered in May of 1994.

For the second cycle of 1998, a team from the Northwest Territories, Saskatchewan, Ontario, Quebec, New Brunswick (francophone), and Nova Scotia (francophone) came together in April 1997 to review the previous assessment and to prepare to re-administer or replicate the assessment. A close analysis of the 1994 assessment statistics and results, advice from statisticians and scorers, and a review of student exemplars informed the discussion.

Overall, educators, students, and members of the public in every province and territory contributed to the evolution of the reading and writing assessment materials. The result was a unique Canadian reading and writing assessment in both official languages.

Curriculum and Criteria

School curricula differ from one part of the country to another, so comparing test data resulting from these diverse curricula is a complex and delicate task. Young Canadians in different jurisdictions, however, do learn many similar skills in reading and writing, mathematics, and science. Throughout the history of SAIP assessments, development teams composed of representatives from various jurisdictions have worked with CMEC staff to consult with all jurisdictions to establish a common framework and set of criteria for each subject area. These were intended to be representative of the commonly accepted knowledge and skills that students should acquire during their elementary and secondary education.

Linking Innovation and Research with Classroom Practices

The role of SAIP has been from its inception to provide educational leadership by producing assessments based on current innovation and the most contemporary research and practice related to student assessment. The goal has been to link innovation and research with familiar and current classroom practices. To this end, the Writing III Consortium developed a new design for the writing assessment. While maintaining the dominant characteristics of the criteria from the previous two assessments, the new design embraces a more contextualized and cross-curricular framework. As well, an additional writing activity was added to introduce the possibility of reviewing general skills related to critical thinking.

Five Levels of Achievement

Achievement criteria for SAIP assessments are described on a five-level scale, representing a continuum of knowledge and skills acquired by students over the span of their elementary and secondary experience. It is important to realize that the same assessment instruments are administered to both age groups (13-year-olds and 16-year-olds) to study the change in student knowledge and skills due to the additional years of instruction. The development teams therefore designed an assessment in which most 13-year-olds would be expected to perform at least at level 2 and most 16-year-olds, at least at level 3. In spite of potential differences in course selection by individual students at secondary school, SAIP assessments should still help to determine whether students attain similar levels of performance at about the same age.

A Program Assessment, Not a Student Achievement Assessment

In the SAIP assessments, the achievement of individual students is not identified, and no attempt is made to relate an individual's achievement to that of other students. The SAIP assessments essentially measure student performance in a subject and reflect this back to each jurisdiction. These assessments do not replace but rather complement individual student assessments, which are the responsibility of teachers, school boards, and ministries of education. The results are reported at the pan-Canadian and jurisdictional levels only.

Assessment Over Time

An important factor to be considered is the impact of changes in curriculum and in teaching practice over time that result from developments in education research and changing public understandings of the role of education in society. Generally, SAIP assessments in all subject areas are designed to retain sufficient elements from one administration to the next to allow longitudinal comparisons of student achievement, while making certain modifications to reflect changes in educational policies and practices. However, considerable caution is necessary in comparing the 2002 writing results with those for both 1994 and 1998 in light of changes to the framework and design of the new instrument (see section below on Conceptual Framework and Criteria). The 1994 and 1998 writing assessments were virtually identical, whereas the topic, scoring criteria, and procedures for 2002 were more precisely defined, with an additional element required. The writing prompt for the assessment writing task was more prescriptive than that of 1998. The 1998 assessment introduced the general theme of *Heroism* and asked students to write about that theme in a form that was comfortable to them. In 2002, students were given a specific real-life environmental dilemma and asked to *generate public awareness about this dilemma*. As a result of the more prescriptive prompt, the criteria described student writing as demonstrating *the elements of writing appropriate to purpose*. As well, at level 2, the descriptive expression *uncertain grasp of the elements of writing* became *uncertain control of the elements of writing*. Experienced educators studied a wide range of actual student work and selected anchor papers for scorers that were valid illustrations of the criteria in light of the demands of the prompt. These changes are significant enough to suggest that any attempt to compare student performances of 2002 with those of 1994 and 1998 must take into account the changes in design and scoring. Therefore, this public report does not directly compare the SAIP Writing III results with past writing assessments.

Nevertheless, specific jurisdictional analysis suggests that a form of comparison can be made by referring to comparative percentage differences in performance by gender, age, and levels between 1994, 1998, and 2002: for example, if 16-year-old girls were 15% stronger than boys at level 3 in 1998, how does this compare with the difference between the same groups at level 3 in 2002?

HARMONIZATION OF ENGLISH AND FRENCH

From the outset, the content instruments used in all SAIP assessments are developed by anglophone and francophone educators working together for the purpose of minimizing any possible linguistic bias. During the development of every aspect of the assessment including materials, administrative handbooks, and scoring procedures, every effort is made to ensure equivalence between both languages. For the SAIP Writing III these efforts included the following:

For the *Student Resource Booklet*, care was taken to ensure equal proportions of French and English source documents as resources. Among the approximately 18 segments,

- eight were from sources previously published in both languages;
- three were translated from French to English and three from English to French;
- four were in English only and four in French only, of which three were poetry and one a brief passage from a newsmagazine, all of them generally equivalent in reading level and theme.

Table 2	
Overview of the SAIP Writing III Assessment (2002)	
Participating jurisdictions	All 10 provinces and 2 territories
Population sampled	13-year-old students and 16-year-old students [The identical assessment was administered to both age groups.]
Number of participating students	23,680 students • 12,708 thirteen-year-olds • 10,972 sixteen-year-olds
Languages in which the test was developed and administered	Both official languages • 17,980 anglophone students • 5,700 francophone students • Provinces with significant sampling populations in both languages received results for both language groups.
Framework	• Writing for a particular purpose to a particular audience • Responding to a brief text
Assessment administration	• Five days or less prior to the writing activity, students spend an hour responding to the theme and discussing resources provided. • All students were given a 2.5-hour period in which to draft, revise, and polish a piece of writing. • All students completed a student questionnaire. • The teacher and principal each completed separate questionnaires.
Results	• Reported for Canada • Reported for jurisdictions • Pan-Canadian expectations set by a broadly representative panel of Canadians
Scoring	• Five levels of achievement in writing • Eight codes for a written response to a brief thinking activity
Reports	• Public report [this report] • *Student Writing: The Canadian Context* [context report] • Technical report

The instrument and school administration manuals were drafted by a bilingual team. The process included feedback about equivalence in language from both anglophone and francophone jurisdictions during the consultation process. As well, all documents were submitted to a formal linguistic revision process.

Care was exercised in the field testing of themes, materials, formats, and sampling procedures so as to ensure equivalency. Final decisions regarding these elements took into account feedback from students and teachers in both language groups.

The scoring was conducted concurrently in both languages. Procedures included the following:

• parallel training of both table leaders and scorers;
• a bilingual committee with responsibility for reviewing all selections of anchor papers to ensure comparability at every level;

- bilingual scoring tables in place, one trained with francophone scorers, the other with the anglophone group. Scorers assigned to these tables scored booklets in both languages, including a sample of identical booklets to track the consistency in scoring practices between language groups.

In April and May 2002, the SAIP Writing Assessment III was administered to a random sample of students drawn from all participating jurisdictions. Approximately 23,700 students made up the total sample — 12,700 thirteen-year-olds and 11,000 sixteen-year-olds. Students completed the assessment in their first official language; about 18,000 students wrote in English and about 5,700 in French. Students in French immersion wrote in English. Detailed breakdowns of the numbers of students assessed in each jurisdiction are presented in the appendix.

The writing assessment involved two sessions.

- The first session, which was approximately one hour long, allowed students to familiarize themselves with the theme of the writing task by first responding to a short text for 20 minutes. This first session was to precede the second session by five days or less. They then discussed a series of brief texts that explored the theme in a resource handbook.
- In the second session students had two and a half hours to fulfil the assigned writing task.

Students who received special accommodations in their regular classrooms were allowed those accommodations for this assessment. For example, students who normally had a scribe to write were permitted a scribe for these assessments. Braille or large-print tests were also provided as needed. Students were given extra time to complete the assessments if, in the judgment of the school-based staff, they required it.

All students sampled were asked to complete a background questionnaire that described some of their reading and writing habits and interests. In addition, the 2002 assessment included questionnaires for both teachers and school administrators. Results will be presented in the document *Student Writing: The Canadian Context* and in the technical report.

DEVELOPMENT OF THE ASSESSMENT DESIGN AND MATERIALS

In order to continue to offer writing assessments based on current innovation and the most contemporary research and practice related to performance assessment, both the revised writing task and the thinking activity were developed after thorough study of current research and literature, and following consultation with the ministries of education.

Foundations and Design

The most current research and practices in writing pedagogy and assessment emphasize that writing assessments should not simply measure whether students can produce a particular text, but whether they can apply knowledge of writing for a specific purpose in a specific context. A writing assessment prompt should not represent writing as an end in itself but rather should propose an activity that allows writers to communicate effectively in some larger "real-life" situations for reasons a classroom community might experience as authentic.

To this end, the writing prompt for this assessment was more prescriptive than that of 1998. The 1998 assessment introduced the general human theme of *Heroism* and asked students to write about that theme in a form that was comfortable for them. In 2002, students were given a specific real-life environmental dilemma and asked to write *to generate public awareness about this dilemma*.

The general theme of the writing task and support materials was *Sharing Living Spaces*. This was intended as a cross-curricular theme linking environmental, scientific, social, and political

information and issues relevant to both classrooms and local communities. It was intended to take the assessment out of the language arts classroom and place it in the larger context of writing needs for learning and living in the broader community.

Current research on writing also suggests that rarely is writing done in isolation. In school and in the workplace, most writing activities involve some "scaffolding" of resources, direction, and consultation. In a time-limited writing task, it is important to provide a knowledge base for writers so that focus can be on quality of writing rather than on the limitations of the knowledge and concrete reference each student brings to the task. To this end, the task design included

- a resource handbook
- a brief reading response activity to stimulate engagement with the theme
- specific time allotted to discussion and reflection with classmates, teachers, and parents using the texts from the resource booklet
- clear instructions with suggestions about ways to approach the task and the forms a student might choose
- a checklist for revision

The Student Resource Booklet

In keeping with "best practices" for the writing process, a student resource handbook was provided for pre-writing activities. The resource booklet

- established and defined the theme;
- offered a wide range of types of text, with a variety of length and reading demands, from simple to complex, including editorials, letters, explanatory articles, personal essays, charts, pictures, short information paragraphs, cartoons, stories, poems, news articles, and quotations;
- followed a carefully planned sequence and development. The materials move from texts that treat the theme in a general fashion to those with specific examples of issues raised and approaches taken. Then the materials focus on a particular animal species faced with "sharing living space" with humans.

Pre-Writing Activities

In the first preparatory session, students began by reading a short narrative text and responding to it in writing. The consortium chose a passage that was short and simple enough in language so as to be accessible to weaker readers, but challenging enough in its thematic intent to be stimulating for both age groups. Students were required to read and respond to this text in a 20-minute period. Given these constraints, a fable proved to be the most effective choice because this genre combines a simple and short narrative with a sophisticated theme. The fable chosen had both a concrete descriptive narrative and a clearly stated moral that provided the student with an abstract issue applicable to other human experiences. Moreover, the moral conveyed by the fable was open to question, thus providing a stimulus for critical reflection.

The prompt itself was carefully crafted to encourage student thinking without directing it. Words and phrases were used to prompt reflection without imposing an academic activity; *think about it* [the text] *carefully, suggest what it means to you, explain your ideas thoughtfully*.

Following this initial writing activity, students were to take part in a discussion of two short texts from the resource booklet. They were encouraged to continue reading the resource materials at home and to discuss them with classmates, with family members, and with other adults prior to the second session. Students were also instructed to bring the resource booklet with them to the second writing session along with any reflections they might have gathered on the notes page provided.

Within five days of the pre-writing session, students reconvened for the main writing task. A specific environmental dilemma was outlined. The writing context was described as a science classroom. The purpose was to provide an environmental agency with a written text that would help *generate public awareness about the dilemma*. Students had two and a half hours in which to draft, revise, and polish their writing.

The Development Process

The pre-pilot proposal

The consortium built a design profile and rationale reflecting the principles of a contextualized prompt with a cross-curricular theme and accompanying resources. Coordinators considered the changes and provided feedback. With approval in principle for a new prompt design, the team prepared the pilot documents.

The pilot process

In both English and French, five different themes, prompt designs, and resource booklets were shaped to fit the criteria and framework recommended in the proposal. The prompts varied in design to test a range of possible approaches within the parameters of the conceptual framework. However, all prompts established a clear context and theme with a problem to be resolved or explored in writing. These were tested through a cross-country sample of 3,000 and 2,000 students respectively in English and French in language arts classrooms with a range of age groups in grades 8 and 11. Brief questionnaires sought feedback from both students and teachers.

Teacher and student comments were examined; however, a higher priority was given to the actual student writing elicited by each prompt. While other factors such as teacher enthusiasm, pre-writing discussion, and the immediate context of the pilot in any one school can all affect student performance, the development team identified five categories within which to judge the success of each prompt:

• commitment or interest level apparent in the writing
• the range of discourse or forms generated with the assumption that a broader range of forms implies greater accessibility and awareness of possibilities for expression
• the quality of writing, assuming that the better prompt allowed students to produce better writing
• the degree to which students seemed to grasp the purpose and context of the writing demand
• equivalences or the degree to which a prompt appeared equally effective in both languages

Qualities of the prompt chosen from the pilot evidence

The student writing from the most effective prompt confirmed current research on writing assessment and prompt design. The prompt was highly contextualized. It referred to a specific dilemma and to a general purpose for writing in response to that dilemma. The need for prior knowledge for content was limited. Specific resources were provided so that students did not need extensive prior knowledge and could focus on the writing activity rather than struggle for references. The resource booklet provided general thematic underpinnings for the writing situation and encouraged pre-writing reflection. Not only was there a clear encouragement and opportunity to use either narrative or analytic forms of expression, such forms were modelled in the resource booklet.

In both languages, students responded well, with imagination and commitment and a clear sense of the specific writing purpose. Some teachers considered the prompt too specific and less interesting because it was more akin to a science class activity than to a language arts exercise. However, a wide range of forms were taken up by students, and many students commented on the fact that they learned some new things from the resource materials.

Consultation and feedback

The selected pilot was polished, and, along with a conceptual framework statement and a set of criteria for assessment, it was submitted to the ministries of education for further feedback and revisions. Ministries of education were consulted at three stages of refinement following the piloting of the design.

A series of external experts, both francophone and anglophone, from across the country were also consulted; these included professors of education, literacy assessment specialists and scholars, and teacher representatives. This validation exercise provided further direction for refinement of the instrument.

Consultation focused on the key differences between the previous writing assessments and Writing III. While the 1998 writing task asked students to respond by writing about a general theme, the 2002 task required students to attend to a specific situation and write for a specific purpose. Students did have to take up the challenge "to generate public awareness about this dilemma." The context was clearly more aligned with science and social studies than language arts, yet the resource booklet included a wide range of literary as well as non-literary, informational and media texts. The issue of Sharing Living Spaces was seen as a very contemporary and public issue. The purpose of the writing was intended to provide a highly authentic motivation to write.