INTRODUCTION

SCHOOL ACHIEVEMENT INDICATORS PROGRAM (SAIP)

Canadians, like citizens of many other countries, want their children to have the best educational preparation possible. Consequently, they are asking how well our educational systems prepare students for lifelong learning and for participation in the global economy.

To help answer this question, ministries[1] of education have participated in a variety of studies since the mid-eighties. At the international level, through the Council of Ministers of Education, Canada (CMEC), Canadian provinces and territories took part in the international Indicators of Educational Systems program of the Organisation for Economic Co-operation and Development (OECD). Individual jurisdictions participated in various achievement studies such as those of the International Assessment of Educational Progress (IAEP) and the International Association for the Evaluation of Educational Achievement (IEA). Typical of international studies are the Third International Mathematics and Science Study (TIMSS), in 1995, and its replication in 1999, and the OECD Programme for International Student Assessment, to be administered in 2000. In addition, in most jurisdictions, ministries undertook or enhanced measures at the jurisdictional level to assess students at different stages of their schooling.

Since all ministers of education wish to bring the highest degree of effectiveness and quality to their systems, they have long recognized a need for collective action to assess these systems. They acknowledge that achievement in school subjects is generally considered to be one worthwhile indicator of the performance of an education system. In particular, the ministers wanted to answer as clearly as possible the question: "How well are our students doing in mathematics, reading and writing, and science?"

In that context, in 1989, CMEC initiated the School Achievement Indicators Program (SAIP). It was a first-ever attempt by all the ministers of education to arrive at a consensus on the elements of a national assessment. 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 to administer the assessments in the spring of each year as follows:

Mathematics	Reading and Writing	Science
1993	1994	1996
1997	1998	1999
2001	*Assessments to be determined in 2002, 2004*	

The first cycle of assessments took place as scheduled, and the reports were published in December 1993, December 1994, and January 1997, respectively. The second cycle has also proceeded as scheduled, with the second science assessment administered in the spring of 1999.

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

Because this is the second science assessment, two questions are asked. In addition to the initial question: "How well have Canadian 13- and 16-year-old students learned science in 1999?" there is also the question: "Has the achievement of Canadian 13- and 16-year-old students in science changed since 1996?"

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. The SAIP assessments should help determine whether these students attain similar levels of performance at about the same age.

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 how well the education system of each jurisdiction is doing in teaching the assessed subjects. They do not replace individual student assessments, which are the responsibility of teachers, school boards, and ministries of education. Similarly, no attempt is made to compare schools or school districts. The results are reported at the Canadian and jurisdictional levels only.

Funding for the SAIP assessments is provided jointly by CMEC, ministries of education, and Human Resources Development Canada.

SCIENCE EDUCATION IN CANADA

The Science Council of Canada, in its 1984 report *Science for Every Student: Educating Canadians for Tomorrow's World,*[2] described the importance to Canada of its citizens acquiring a good working knowledge of science concepts and the inquiry skills that allow the application of these concepts to everyday life in the world around them. The report endorsed the concept of a science education program for each of Canada's students, in all regions and provinces, regardless of ability and interest that would

- develop citizens able to participate fully in the political and social choices facing a technological society
- train those with special interest in science and technology fields for further study
- provide an appropriate preparation for the modern work world
- stimulate intellectual and moral growth to help students develop into rational, autonomous individuals[3]

Since the release of this influential report, curriculum development in Canada and in other countries has emphasized the importance of developing a scientifically literate population, while at the same time providing for those students with special aptitudes and interest in these fields, opportunities to grow in a challenging learning environment. The evolution of a significant role for Science–Technology–Society–Environment (STSE) in emerging curriculum is a strong indication of the influence of this report and others like it.

Typical of this development in Canada are two recent publications. An important pan-Canadian initiative was the 1997 release of the *Common Framework of Science Learning Outcomes*,[4] a product of the CMEC Pan-Canadian Protocol for Collaboration on School Curriculum. This document is intended as a tool for the developers of future science curricula. An example of the interest taken in science education by a nongovernmental organization is the publication of *Science Literacy and the World of*

---

[2] Science Council of Canada. *Science For Every Student: Educating Canadians for Tomorrow's World*. Report 36. Ottawa: Science Council of Canada, 1984.
[3] Ibid.
[4] Council of Ministers of Education, Canada. *Common Framework of Science Learning Outcomes*. Toronto: Council of Ministers of Education, Canada, 1997.

Work[5] by the Conference Board of Canada. This pamphlet describes "scientific, technological and mathematical competencies for an innovative, productive, and competitive workplace."

The SAIP Science Assessment Framework and Criteria reflects the intent of many of these recent initiatives. While the understanding of the process of teaching and learning about science is continually being refined, the Framework and Criteria used in 1999 is identical to that used in 1996. This is to allow the comparability of results between the two assessments — an important feature of SAIP.

LEARNING ABOUT SCIENCE

Science is more than a body of complex and abstract knowledge and theories about the universe. It is also more than a set of processes that guide scientific inquiry and discovery. While both of these images of science are important to the working scientist, for effective learning, science must relate to the everyday life of students and must engage them in the process of learning about the world around them. All students learn most effectively about their world by guided, direct observation and hands-on experiences that allow them to gain knowledge and acquire skills in an environment in which they find application to their daily lives.

> By gaining lots of experience *doing* science, becoming more sophisticated in conducting investigations, and explaining their findings, students will accumulate a set of concrete experiences on which they can draw to *reflect* on the process. At the same time conclusions presented to students … about how scientists explain phenomena should … be augmented by information on how the science community arrived at these conclusions.[6]

THE ASSESSMENT OF SCIENTIFIC LITERACY

For many students, the SAIP Science Assessment was a new and different testing experience. Rather than a test that emphasizes the simple recall of information, students encountered an assessment that asked them to relate their understanding of science to real-life situations that were familiar to them.

In the written component of this assessment, students' knowledge of science concepts and their application to society around them, as well as their understanding of the nature of science, were measured by responses to multiple-choice and short, written-response questions. For those who participated in this part of the assessment, the questions were presented in groups within simple and common scenarios that required the application of knowledge to situations familiar to young people.

While the attainment of science inquiry skills is universally acknowledged to be an essential aspect of science education, the assessment of achievement in this area, particularly on a large scale, often has been seen as difficult, if not impossible. The SAIP Science Assessment attempted to achieve this goal by conducting a hands-on practical task assessment, challenging students to apply the science inquiry and problem-solving skills found in the *Science Assessment Framework and Criteria*[7] to simple hands-on tasks. For each of the seven tasks, students who participated in this part of the assessment were asked to respond to a series of questions that assessed their level of understanding of specific science skills.

---

[5] The Conference Board of Canada. *Science Literacy and the World of Work*. Ottawa: The Conference Board of Canada, 1996.

[6] American Association for the Advancement of Science. *Benchmarks for Science Literacy*. New York: Oxford University Press, 1993.

[7] See page 9.

OVERVIEW OF THE 1999 SCIENCE ASSESSMENT

In April and May 1999, the science assessment was administered to a random sample of students drawn from all provinces and territories. Approximately 31,000 students made up the total sample — 16,000 thirteen-year-olds and 15,000 sixteen-year-olds. About 22,500 students completed the science assessment in English, and 8,500 in French. For some provinces and territories, where the total number of students was small, the whole age group population was selected. Detailed breakdowns of the numbers of students assessed in each jurisdiction are presented in the appendix of this report.

As in other SAIP assessments, students in both age groups responded to the same assessment instruments. Participating students were asked to complete one of two components. A written assessment focussed on their knowledge of science concepts, the nature of science, and the relationship of science to technology and societal issues. A practical task assessment focussed on science inquiry skills by presenting practical problems in a hands-on environment. For the written assessment, a representative sample from each jurisdiction was drawn of sufficient size to allow reporting at the national level as well as the jurisdictional level. A national sample was drawn by CMEC for the practical task assessment. In addition, three jurisdictions — Saskatchewan, Ontario, and Quebec — took the option of over-sampling all or part of their sample populations to allow reporting data at a provincial level.

As in the past, to assist with the interpretation of outcomes for the SAIP Science Assessment 1996, CMEC convened a pan-Canadian panel of educators and non-educators, each of whom attended one of the three sessions held in Western, Central, and Atlantic Canada in October–November 1999. A collaborative, two-stage process was used to define pan-Canadian expectations for student performance in both the written and practical components of the assessment. Details of the process and the results of their deliberations can be found under Pan-Canadian Expectations, page 27.

COMPARABILITY OF THE 1996 AND 1999 ASSESSMENTS

An important reason for conducting this assessment, only three years after the first science assessment, was to ask the question: "Has the achievement of Canadian 13- and 16-year-old students in science changed since 1996?" A primary concern of the 1999 development team was to ensure that changes to assessment instruments and scoring procedures were kept to a minimum. The same framework and criteria were used to assess 1999 student work. Scoring procedures and conditions were replicated as much as possible from information provided by the previous team.

Changes to the written assessment instruments were restricted and minimal:

- corrections to typographical or minor linguistic matters
- standardizing formatting for all questions and ensuring a consistent format for both French and English
- minor wording changes for clarity in a very few places
- including all questions in one booklet with three colour-coded sections, rather than three separate booklets as in 1996

Changes to the practical task assessment were somewhat more significant. Since one of the tasks had been compromised through inclusion in the 1996 public report, it was replaced. In addition, minor modifications were made to two others to remove some unnecessarily confusing wording.

For 1999, the scoring of the practical task booklets was enhanced by two factors:

- a greater number of experienced scorers
- more clearly defined scoring criteria (not different criteria, but more clearly defined)

A combination of these factors enabled the professional educators who were the scorers to exercise their professional judgment in a more consistent manner. The effects of this enhanced scoring process are expected to have improved the consistency of scoring and thus the confidence that one can have in the results.

All changes were subject to fresh field trials in the schools of the consortium jurisdictions to ensure the appropriateness of the changed instruments.

In all other ways, the assessment materials were the same. For the written assessment, a placement test was administered, in which students were asked to answer a preliminary set of twelve level 3 questions to assist in assigning the levels of remaining questions. Administration procedures for both the written and practical task components were essentially the same in 1999 as in 1996.

In the sampling procedure, student selection was modified slightly from the 1996 assessment. In 1999, students were selected without any exclusions, while in 1996, students could be excluded before the final sample was drawn. In 1999, school administrators, together with school staff could consider that a student had very limited abilities in science and that it would serve no purpose to have the student write the assessment. If the student could not make a reasonable attempt at answering any of the level 1 questions included in the *Information Bulletin for Schools*, the school could exempt the student and designate him or her as below level 1. It is therefore likely that more students were included in the 1999 sample that would not meet the criteria for level 1.

A second source of comparability from 1996 included the involvement of a number of scoring leaders and scorers from the 1996 sessions in the 1999 scoring sessions for practical task booklets. This assisted in establishing similar scoring communities with similar contexts for scoring.