BASIS OF THE ASSESSMENT

ASSUMPTIONS AND LIMITATIONS OF THE ASSESSMENT

Although the content of this assessment was consistent with that of science curricula across Canada, it could not be comprehensive enough to include everything that appears in every science program. It is as much an assessment of scientific literacy as a science test in the usual meaning of the word. The assessment focussed on knowledge and skills that can be measured in paper-and-pencil testing and on practical tasks. The teamwork or cooperative problem-solving approach, often used in solving scientific problems, was not evaluated in this assessment.

In both assessments, scoring was based upon a comparison of students' responses to the criteria in the Science Assessment Framework and Criteria upon which the items were based. For the written assessment, recent faculty of education graduates, using a template of acceptable responses, scored the extended-response (written) questions. In the case of the practical task assessment, experienced science educators were trained to compare student responses to exemplars chosen from actual student papers by the development team. A number of scoring leaders and scorers returned from the 1996 administration. This ensured increased consistency in the scoring process.

DEVELOPMENT OF THE 1996 ASSESSMENT MATERIALS AND THEIR REVISION FOR 1999

The 1996 Assessment

The development of the components of the 1996 SAIP Science Assessment began in the fall of 1993 when CMEC asked the ministries of education in Alberta, Saskatchewan, Ontario, and New Brunswick (francophone) to form a consortium of subject and assessment specialists. These specialists were asked to develop science material that would describe and assess five levels of achievement for 13- and 16-year-olds. The consortium worked in cooperation with other ministries of education.

Provincial curricular materials present science as a continuum of learnings from elementary through to the end of secondary school. Four areas within these science learnings form the framework for this assessment:

- knowledge and concepts of science
- nature of science
- relationship of science to technology and societal issues
- science inquiry skills

Criteria for the assessment were drafted to reflect the breadth of what students in Canadian schools are expected to know and be able to do with respect to these four areas. In keeping with the current emphasis on conceptual understanding of science, points of progress along the continuum were organized to represent five levels of progress.

As the Assessment Framework and Criteria evolved, each ministry of education reviewed draft proposals in the context of its curriculum and according to its own consultation procedures. Classroom teachers and professional groups also reviewed the criteria and proposed assessment framework. Their concerns and suggestions directed subsequent revisions of the criteria and the assessment

design. Student evaluation and curriculum specialists from the universities, science experts, and representatives from nongovernmental organizations also reviewed the criteria. Teachers from across Canada developed questions and tasks for the assessment during the summer of 1994. Each ministry was then asked to carry out a curriculum and bias review of this material.

A first informal field test of the questions was carried out in the fall of 1994 in a limited number of classrooms in the four consortium provinces. In the spring of 1995, the assessment materials, including twice the number of items needed for the final test, were fully field-tested in all the provinces. Comments by teachers whose students had field-tested the instruments were very useful in the revision process. The developers also considered students' comments about the questions, the tasks, and the administrative procedure. Field-test scorers' comments and test results confirmed the appropriateness and range of difficulty of the questions, tasks, instructions, and administrative procedures. Particularly in the case of the practical task assessment, the deliberations at the marking session also confirmed the effectiveness of the criteria and the procedures for scoring in order that students would be placed at the appropriate skill level.

The 1999 Assessment

In April 1998 a team from Saskatchewan, Ontario, Quebec, Nova Scotia (francophone), and Newfoundland and Labrador came together to review the assessments and prepare them for re-administration. A close analysis of all 1996 assessment statistics and results, advice from statisticians and scorers, and a review of student exemplars informed the discussion. As described earlier,[8] changes to assessment instruments and scoring procedures were kept to a minimum. The same Framework and Criteria was used to assess 1999 student work. Scoring procedures and conditions as well as administration procedures were replicated as much as possible from documentation and information provided by the previous team.

In all of its work, the 1999 consortium team strove to make the second cycle of the assessment comparable to the 1996 assessment. Attention was paid to this factor at all levels — instrumentation, administration, scoring, data collection and analysis, and reporting.

ADMINISTRATION OF THE ASSESSMENT

Written Assessment

All students writing this assessment began by doing 12 questions at level 3. On the basis of their scores on those 12 questions, students were directed to a subsequent particular set of colour-coded pages in their test booklet. Each set of questions contained 66 items that covered a different combination of achievement levels. Section B covered levels 1, 2, and 3. Section C covered levels 3, 4, and 5, level 5 being the highest. The 66 questions in each section were a combination of multiple-choice and written-response questions. All students, regardless of which set of items they progressed to, wrote an identical set of 26 level 3 questions, 12 from the placement test and 14 repeated in each of parts B and C.

Practical Tasks

Specially trained external test administrators brought the hands-on testing materials to the sample schools and administered the assessment to the selected students. Students participating in the science inquiry skills assessment performed seven tasks that required them to generate and analyse their own data by applying science inquiry skills to questions of a scientific, technological, and/or societal nature.

[8] See page 4, "Comparability of the 1996 and 1999 Assessments."

The written assessment was scored in Sudbury, in June 1999 and the practical task assessment was scored in Montreal, during July and August. Data processing took place in Quebec City; statistical analysis was carried out in Vancouver. A consultant prepared drafts of the report for approval by the CMEC Secretariat, in cooperation with the Science Assessment Administration Management Team and the Report Development Group.

Contextual Data

Questions regarding opportunities students have had to learn science, some of their attitudes toward science and other demographic information were gathered in a student questionnaire.

For the 1999 assessment, in order to collect a broader selection of contextual information, school principals completed a school questionnaire, and science teachers were asked to complete a science teacher questionnaire.

COMPARISONS BETWEEN LANGUAGE GROUPS

From the outset, the instruments used in the science assessment were developed by English- as well as by French-speaking educators working together for the purpose of eliminating any possible linguistic bias. Whether they wrote in French or in English, the students responded to the same questions and executed the same tasks. Consequently, the statistical results presented for each language group in this report can be compared with reasonable confidence.