University of Alberta

THE SCHOOL ACHIEVEMENT INDICATORS PROGRAM

1988 - 1996

BY

KAREN SLEVINSKY 

A thesis submitted to the Faculty of Graduate Studies and Research in partial fulfillment of the requirements for the degree of Master of Education

Department of Educational Administration

Edmonton, Alberta

Fall 1996

ISBN 0-612-18222-3

University of Alberta

Library Release Form

Name of Author: Karen Slevinsky

Title of Thesis: The School Achievement Indicators Program 1988 - 1996

Degree: Master of Education

Year this Degree Granted: 1996

K Slevinsky

Permanent Address:

32 Blue Quill Crescent
Edmonton, Alberta
T6J 6C4

Date: October 2, 1996

June 17, 1996

Ms. Dianne Pennock
National Coordinator SAIP
Council of Ministers of Education, Canada
252 Bloor West, Suite 5-200,
Toronto, Canada
M5S 1V5

Dear Ms. Pennock,

I am a graduate student at the University of Alberta in the Department of Educational Policy Studies. I was a member of the SAIP Science Development Team. I am nearing the completion of a thesis entitled The School Achievement Indicators Program 1988 - 1996, An Historical Survey and Analysis.

To enhance my thesis and increase its usefulness, I would like to include a copy of the Memorandum of Understanding for the School Achievement Indicators Program, dated May 20, 1994. I am requesting permission to reprint this MOU and add it to my thesis as an appendix.

Thank you for your attention to this matter and also for the assistance provided to me by the CMEC Secretariat in other related matters. I look forward to your reply.

Yours truly,

Karen Slevinsky

Karen Slevinsky
32 Blue Quill Crescent
Edmonton, Alberta
T6J 6C4

Aug. 14/1996
Approved

University of Alberta

Faculty of Graduate Studies and Research

The undersigned certify that they have read, and recommend to the Faculty of Graduate Studies and Research for acceptance, a thesis entitled **The School Achievement Indicators Program 1988 - 1996** submitted by Karen Slevinsky in partial fulfillment of the requirements for the degree of Master of Education.

M. Haughey

Dr. Margaret Haughey
(Supervisor)

Dr. Wytze Brouwer

Dr. J.L. Da Costa

Date: September 24, 1996

Dedication

To Richard and Mikaël

for their patience, understanding, and above all, their encouragement

and support.

Abstract

The purpose of this study was to trace and describe the chronological development of the national School Achievement Indicators Program (SAIP) from 1988 until 1996. The context for the genesis of SAIP and its initial development is described. The SAIP involves testing a sample of 13 and 16-year-old students in Canada to determine the level of achievement of similar students in mathematics (1993 & 1997), reading and writing (1994 & 1998), and science (1996 & 1999). The specific SAIP program examined is the science assessment. An explanation of how the SAIP science assessment instruments were developed including their framework, criteria, item-writing, field-testing and validation, statistical test and sample design, and administration is provided. Emerging trends and issues arising out of the establishment of this large-scale indicators program are considered and specific implications of the SAIP for education in Canada and implications for the next science assessment are discussed.

Acknowledgements

I wish to acknowledge my sincere appreciation to all who have made this study possible.

A special thanks to the nine key informants who participated in this study. The time you all devoted to this study is very much appreciated.

A special thanks to the staff of Alberta Education, especially Phill Campbell. Your support and flexibility enabled me to complete the task.

A special thanks to the personnel at Alberta Education and at the secretariat of the Council of Ministers of Education, Canada, who assisted by providing resources and information. Specifically, I thank Dianne Pennock, of the Council of Ministers of Education, Canada, who was especially helpful and accommodating.

A special thanks to Léo Laroche of the Centre de traitment des données du PIRS (SAIP Data Processing Centre), to David Wasserman of Alberta Education, and Richard Slevinsky, my husband, who took the time to read draft versions of the study and provide critical feedback.

A special thanks to Nadia Hochachka, a personal friend at Alberta Education, who offered support as I worked on my study.

And, a very special thanks to my advisor, Dr. Margaret Haughey. The time she took asking the 'hard' and 'pointed' questions and providing formative and critical feedback helped make this a better study. For this, I will always be grateful.

TABLE OF CONTENTS

CHAPTER

List of Tables

List of Figures

Chapter I

INTRODUCTION AND OVERVIEW OF THE STUDY

Introduction

In the spring of 1996, a large random sample of 13 and 16-year-olds wrote either the written or the practical component of the School Achievement Indicators Program (SAIP) science assessment. The SAIP science assessment was a quantitative study based on a predetermined framework and set of criteria. Its instruments were criterion referenced tests. These instruments were used to determine the level of achievement and the extent of growth of the students writing them and to predict the general level of science competency for Canadian 13 and 16-year-olds. This assessment was the third in the series of the Council of Ministers of Education, Canada (CMEC) assessment initiative. The SAIP mathematics assessment was held in 1993 and the reading and writing assessment in 1994. The history of the establishment of the SAIP and the history of the development of the SAIP science assessment is the focus of this research study.

In Canada, the SAIP assessments were preceded by other similar national and international assessments. In October of 1989, Statistics Canada conducted a survey of literacy skills used in the daily living of adults aged 16 to 69 years. The literacy skills included in this survey were reading, writing, and numeracy. Canadians have participated in various American and international studies such as the Second International Science Study (1985), the International Assessment of Educational Progress science and mathematics assessment (1991), and the Stevenson Study of Mathematics Achievement (1993). Students in Canada also participated in the Third International Mathematics and

Science Study (1995) (TIMSS). The results of these studies provided some information to Canadian educators, decision-makers, and the public but this information was not complete since not every province participated in any one study. The CMEC assessment initiative provides information about student achievement against an agreed upon set of criteria in Canada and the student achievement of each province and territory. Cross provincial comparisons can be used to infer the effect of the varied curricula and perhaps provide evidence for the need for national standards in education.

Statement of the Problem

Since Canada has few national initiatives in education and in assessment in particular, it is important to document the process used to develop such a program. Therefore, the purpose of this study is to provide an historical description of the SAIP. Its establishment, development, and implementation have been researched and analyzed to provide an account of its activities from its beginnings in 1988 to the completion of the three subject area assessments in 1996. Although the mathematics, reading and writing, and science assessments are considered to be the implementation of the SAIP, this study focuses on the description of the process of development of only one of the assessments; the science assessment of 1996. The following questions guided the study.

Research Questions

1. What was the rationale for the genesis and establishment of the School Achievement Indicators Program?
2. How did the SAIP develop? Who provided the direction and leadership?

3. How was the science assessment program conceived and implemented?
4. How were the framework, criteria, and the instruments of the science assessment designed, constructed, and administered?
5. What issues emerged in the implementation of the assessment programs and how were these resolved?
6. What are the implications for the next SAIP science assessment in 1999?
7. What themes emerged as the SAIP was established and as the SAIP science assessment was developed and implemented?
8. What are the implications of the SAIP for education in Canada?

Significance of the Study

This study has both practical and theoretical significance. Since it is the intention of CMEC to repeat the assessments three[1] years after their inception and the science assessment is one component of the SAIP, this study provides five practical contributions.

1. It provides historical information for future SAIP instrument development teams, especially the SAIP Science Development Team for the 1999 science assessment.
2. It identifies the role that the components of the SAIP play in the future of the education of Canadian students.
3. It illustrates the process of establishing and implementing a large-scale project.
4. It documents the issues involved in developing such a large-scale national project.

[1]Since the mathematics assessment was administered in 1993 and will be readministered in 1997 four years will have actually elapsed. Similarly four years will have elapsed for the reading and writing assessment.

5. Quantitative studies, such as the SAIP science assessment, are considered to be based on objective documentation of relevant phenomena. These phenomena are the specific criteria, similar to curricular objectives, upon which this assessment is based. The description of the development, implementation, scoring, and reporting of the science assessment should make it apparent that a large body of extant quantitative data, generated as a by-product of the SAIP science assessment, exists. This study's analysis of the development process of the SAIP should make its qualitative aspects apparent. For this reason, this study makes a contribution to theory. It should also encourage a reader to pose similar questions of any similar research study. In this way, its questions should be transferable to assessment initiatives emerging out of a similar context.

Overview of the Study

The study begins with a review of the pertinent literature, specifically international and provincial indicator projects that were concurrent with the SAIP. Also, readings that explained similar science assessments or science assessment strategies are summarized. The chapter that follows the literature review provides an outline and description of the methods used in this qualitative, descriptive, historical study. Next, chapter four presents the data gathered on the CMEC, the SAIP, and the role of the provinces and CMEC committees in the establishment, the development and the implementation of the SAIP and the SAIP science assessment. The final chapter, chapter five, provides a summary of the findings and discusses the implications of the research study.

Chapter II

EDUCATION INDICATORS: A REVIEW OF THE LITERATURE

Introduction

Based on the examination of some of the major education indicator programs and instruments in the literature, this chapter attempts to capture what others view indicators, indicator systems and indicator instruments to be. Specific indicators programs and assessment strategies have been chosen for examination: the international Indicators of Education Systems; one provincial indicators system, Alberta's Educational Quality Indicators; and a few indicator instruments for determining achievement in science.

Other nations like the United Kingdom have national education monitoring[2] devices but these do not fit the description of an indicators system. These national surveys provide information about one indicator, student performance or achievement, hence they do not necessarily fit the category of an indicators system. Every jurisdiction in Canada has an education indicators program, but the literature refers to the Alberta indicators system most frequently, hence Alberta's Educational Quality Indicators project was examined.

The literature covering the science assessment projects administered in the nineties will be shared. These include the International Assessment of Educational Progress, mathematics and science, and geography (13-year-olds) assessment in 1991 and the International Association for the Evaluation of Educational Achievement TIMSS assessment in 1995. Their testing strategy will be analyzed so as to make similarities and differences to the SAIP science assessment evident. Each of these aspects (international, provincial, and science assessments) of educational indicators will be summarized in order to establish what indicators, indicator systems, and indicator instruments are. This chapter

---

[2]England and Wales administer the Assessment of Performance Unit (APU) while Scotland administers the Assessment of Achievement Programme (AAP).

serves to allow comparisons between the SAIP and the international Indicators of Education Systems, the Educational Quality Indicators, and between SAIP science assessment and other science assessments. The review of the literature illustrates the similar emergence of indicators systems internationally and provincially. It also illustrates the similarities in science assessment strategies among different science assessments.

International Indicators of Education Systems

The international Indicators of Education Systems Project is a system designed to produce comparable international information on various aspects of education. These international education indicators were developed by the Centre for Educational Research and Innovation (CERI) under the direction of the Organisation for Economic Co-operation and Development (OECD). The OECD is an organization of 25 industrialised countries which work together to achieve a high, sustainable economic growth and standard of living. They cooperate on economic development and expansion of world trade practices. The main objectives of the OECD's CERI are the promotion of research activities in education, the promotion of evaluation in educational systems and the promotion of cooperation of member nations in educational research.

During the same years that the CMEC worked on establishing the SAIP, from 1988 to 1989, the CERI studied the feasibility of developing and constructing comparable indicators on education. Various members of the CMEC, notably Alberta[3] and Québec, were involved in this feasibility study. The Centre was interested in determining the quality of education in the OECD countries and in ways to improve it. It proved to be a desirable and feasible initiative and from 1989 to 1991, consensus was reached on the definitions and the selection of indicators and the scope of an indicators system. In 1991, the member nations decided to take action and gather data on some of the agreed upon indicators.

---

[3]In September of 1989, Dr. Reno Bosetti, Deputy Minister of Education, Alberta, Canada, presented a paper on the Canadian experience with establishing an indicators system. This was at the OECD's CERI General Assembly on the International Educational Indicators in Semmering, Austria.

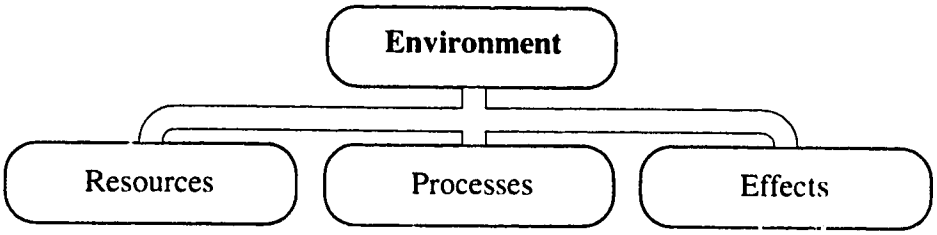
According to the CERI (1992, 1994, & 1995a), education and training frequently appear as the key factors likely to affect the long term competitiveness and productivity of individuals and hence nations. Indicator systems can offer unambiguous information about the functioning of education systems. CERI believes that globalization, which will end educational isolationism, is the key to solving economic, political, environmental and social issues and hence, cooperation and collaboration among member nations is necessary. An international indicators system can demonstrate this. Hence information gathered by indicator systems can be used by educators and policy-makers to monitor, plan, manage, make policies and improve education at the national level. They represent an accountability system, a system that allows practices and polices to be evaluated and revised since indicators capture the attention of educators and policy-makers.

According to the OECD's CERI (1992), education indicators are statistics which provide information about the behaviour and or performance of an education system. Indicators focus on specific aspects of education and they reflect the performance of that part of education. An education statistic is an indicator if it has a reference point against which it can be judged. This reference point is commonly referred to as a standard. Indicators must provide useful, relevant information about the system's performance in achieving desired education conditions or results or about the system's features linked with desired education conditions or results. A single indicator cannot tell everything about an education system. However, a collection of indicators serves to provide a profile of an education system.

The OECD's CERI (1992 & 1994) suggests that since a single indicator is unidimensional, it becomes necessary to build a set of indicators. First, the indicators must be chosen logically. They must represent the components of a particular education system. Second, the selected indicators must be linked. A model or framework describing how an education system works helps determine the linkages among indicators. Choosing the model was difficult for the OECD's CERI since there are differences among the education

systems of its member nations. However, fundamental similarities do exist and a model that acknowledged the environment for learning, resources available, processes, and effects or outcomes of education, emerged. (See Figure 2-1.) A coherent set of indicators, then, is an indicators system.

Figure 2-1 The OECD framework or model of an education system



Source: The OECD International Education Indicators: A Framework for Analysis, OECD's CERI, Paris, 1992, p.19.

The indicators selected by the OECD's CERI (1992, 1994, 1995a, & 1995b) are grouped according to this model. The contextual factors in the environment form one set of indicators. Contextual factors are social, economic, demographic and cultural. Some of these, such as the economic, demographic, and social factors (public satisfaction with and expectations of the education system) can be used as indicators. Information gathered on fiscal and human resources form a group of resource indicators. Processes generally reflect the effectiveness on the education system. Indicators in this category include teacher pupil ratio, hours of instruction, time on task, locus of decision-making, and others that provide information about the features of the school and school system. The effects category considers the results of education. Effects include student, system and labour market outcomes. Some common effects indicators are student achievement, secondary and postsecondary graduation rates, employment patterns of school leavers, earnings and level of education, and unemployment and level of education. As of 1995, the OECD's CERI international Indicators of Education Systems Project had gathered data on a set of 49

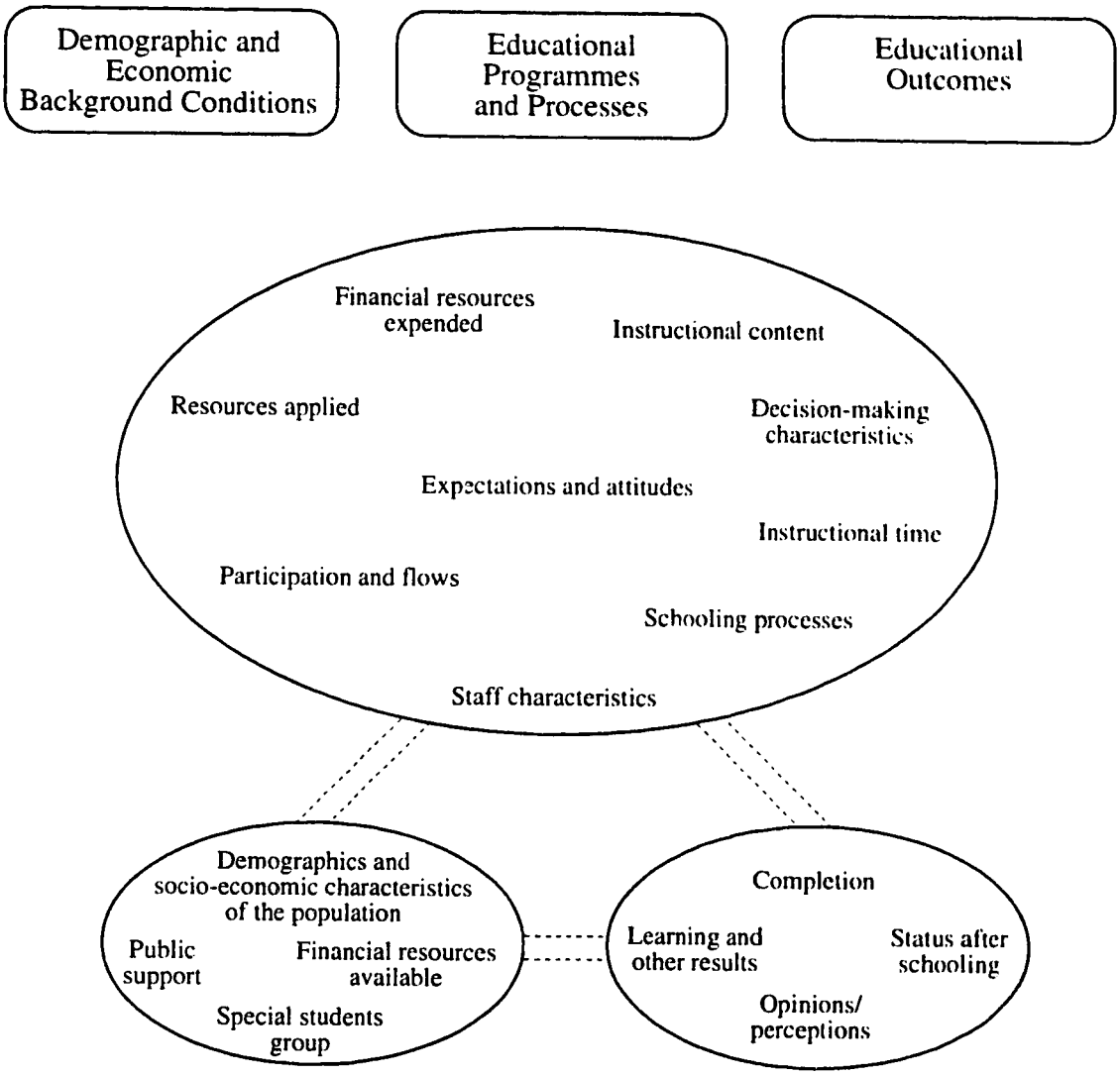
international education indicators, each relating in some way to the underlying model of an education system.

The individual indicators must show some linkage to each other. According to the OECD (1992, 1994, 1995a, & 1995b), sets of indicators should be built on causal relationships among the context, the resources, the processes and the effects of the education system under consideration. Figure 2-2 illustrates the "ideal" set of international indicators for the international Indicators of Education Systems. A comprehensive, yet parsimonious set of indicators makes up an indicators system or program.

Since 1992, the OECD's CERI has collected data, processed it and calculated some of the proposed indicators. These indicators were compiled and published in a series of three editions of Education at a Glance, 1992, 1993, and 1995. A fourth edition is expected to be published in 1996. Each publication is more comprehensive than the last. As more data become available or are collected by the CERI, more indicators are reported. In the 1995 edition of Education at a Glance, each indicator is accompanied by the policy issues surrounding it, a summary of the key results, a short description and interpretation and the definition that the CERI used to define it. For example, for the process indicator, Respect for Secondary Teachers, the policy issue is that this indicator points to support or lack of support for the teaching profession. Of the eleven countries participating in this indicator, the key results (OECD, 1995b, p 59) show that 57.6% of the respondents thought that secondary teachers were "very and or fairly respected." The data charts are described and include the statement that this indicator refers to the proportion of the population responding to the question "In your opinion how respected are secondary teachers as a profession?"

the text continues

Figure 2-2 "Ideal" set of OECD international indicators of education



Source: Indicators of Education Systems Measuring the Quality of Schools, OECD'sCERI, Paris, 1995a p. 20

The OECD's CERI international Indicators of Education Systems Project is an ongoing, evolving project on the collecting and reporting on an ever increasing set of international indicators. The rapid growth of the project is but one indication of the importance member nations, such as Canada, give to this topic.

Provincial Education Indicators

By 1995, five provinces in Canada had developed and implemented some sort of indicators system: British Columbia, Alberta, Ontario, Québec and Newfoundland. Currently every province collects data on some features of its educational system and analyzes some of these indicators. The Educational Quality Indicators was a three year project carried out in Alberta, initiated in 1989[4] and concluded in 1992. The Education Quality Indicators was one of several educational reform efforts to focus on results-based education in Alberta. Similar to the CERI international indicators project, the Alberta Education Quality Indicators was based on a collaborative model and was the initial step in the development of a provincial indicators system (McEwen, 1995a &b).

The Educational Quality Indicators project was a joint initiative that coordinated the development and implementation of ten indicator systems projects with 12 participating school jurisdictions. The Educational Quality Indicators project established three criteria for a successful indicator system: a framework, student outcomes, and points of reference. Each participating jurisdiction was responsible for the development of a conceptual framework, and for an implementation plan for the indicator systems project that looked at student outcomes at various points of reference including time, groups, and targets. The Educational Quality Indicators project administrators recommended that indicators be assessed periodically and regularly, that they should include results from local and provincial groups, and that they should be compared against a standard (Alberta Education, 1993).

Each school jurisdiction developed its own set of indicators based on these three criteria. They developed methods to collect, process, and interpret data, and reported their findings to their own educational community and to the provincial Education Quality

---

[4]In November 1987, the OECD held an international conference on education indicators in Washington, D.C. Although representatives from Canada, a member nation attended, it was noted that government officials from Alberta Education were part of this Canadian team. This conference was one of the earliest major conferences on education indicators.

Indicators group so as to inform participants of their education systems. The Educational Quality Indicators project established a three year phase-in plan. Year 1 was used to develop the indicators project. Year 2 was used to field test the strategy, methods, data collection and interpretation. In year 3, the indicators system was implemented, data were gathered and interpretations were made. Then, the information generated from the indicators system was reported to local and provincial participants.

The 10 projects of the Educational Quality Indicators initiative deal with three aspects of education: system accountability, teaching and learning, and alternative student assessment practices each with its own framework. Four school jurisdictions each developed a set of indicators that provided information about their specific school system. Generally, indicators were collected about finances, quality of instruction, school climate, and student achievement. Two school jurisdictions developed sets of indicators on teaching and learning. Indicators included information on student and teacher behaviour, student achievement, and formative and summative evaluation of teachers. Six school jurisdictions developed indicators of quality student performance, or alternative student assessment practices. Indicators in this aspect of education included exemplars of performance, responsible student social behavior, and portfolio assessment. For example, in the report of the Spirit River School Division No. 46's Educational Quality Indicators' project on teacher effectiveness, 26 indicators of quality instruction were identified. This set of indicators was used to assess teacher effectiveness and was a part of formative assessment of teacher performance. Over a two year period, data were collected on each indicator using five-point rating scales, achievement test and diploma examination results, and by using various other measures. Teachers, administrators, students, and parents were surveyed. Teachers, parents, and students reported agreement with and support for the use of the 26 identified indicators as indicators for teacher effectiveness. Teacher support of this effectiveness study was critical to its success. Spirit River School Division No. 46 recommends the involvement of teachers in the establishment of the set of indicators and has made a

commitment to continue this project. Specific descriptions such as this one make the final report on the Educational Quality Indicators initiative (Alberta Education, 1993) very useful.[5]

The indicator systems developed by the school jurisdictions were based on a general model of indicator systems and tested in practice; the development and the field test phases allowed for this. The final set of indicators provided meaningful information for each school jurisdiction. At the end of the three year Educational Quality Indicators initiative, each school jurisdiction resolved to continue with an indicators system. The indicators school jurisdictions found meaningful were used to develop a set of indicators for a provincial indicators system, as this was the main purpose of the Educational Quality Indicators initiative.

Comparison of the international and the Alberta indicators systems

Both the Alberta initiative and the CERI international Indicators of Education Systems Project reflected the same understandings of what education indicators are and that indicator systems are multidimensional, interrelated and based on a framework of what an education system is believed to be. Like the nations in the international Indicators of Education Systems Project, each participating school jurisdiction built its indicators system on a conceptual framework which reflected the context of its jurisdiction. Unlike the international Indicators of Education Systems Project, no one generic conceptual framework for the province was developed; Alberta's three criteria for a successful indicators system served this purpose. The Educational Quality Indicators and the international Indicators of Education Systems projects are based on the principle of developing a set of indicators that reflected student outcomes. In the Educational Quality Indicators project, these were identified by the participating jurisdiction whereas in the

[5]At the 1994 meeting of the American Educational Research Association in New Orleans, *Achieving Quality, Final Report of the Educational Quality Indicators Initiative* won the Outstanding Publication Award for Policy Studies.

international Indicators of Education Systems Project a concept map of what these indicators are was established for all participating countries. (See Figure 2-2.) The strategy for the development and implementation of the Alberta provincial indicators system is similar to the strategy used by the CERI in the international Indicators of Education Systems Project: consensus and collaboration. Each country participated in the CERI project by submitting whatever educational data it had available or by voluntarily participating in surveys designed especially for the international Indicators of Education Systems. However, Alberta's school jurisdictions submitted their reports to the government of Alberta but did not provide data on their indicators. Both projects had a similar purpose: to improve education by providing data to make improved decisions.

Science Indicators and Assessment Strategy

It is possible to establish indicator systems for an individual discipline such as science education. These indicators would provide information about the state of science education in the school system. Achievement in science would be one indicator in an indicator system that monitors science education. This section will focus on one indicator system and on two different tests which may be used to determine achievement in science. The measurement of the achievement in science of students in jurisdictions, in nations and internationally will be the specific focus of this research. The International Association for the Evaluation of Educational Achievement's TIMSS, the second International Assessment of Educational Progress, and the Assessment of Achievement Programme science assessment will be summarized. The type of assessment instruments and the effect these indicators have on science instruction will be examined.

The International Association for the Evaluation of Educational Achievement's TIMSS

The International Association for the Evaluation of Educational Achievement is an organization that has a long history of conducting comparative studies in mathematics and

science, as well as in other disciplines. A consortium of educational research institutions in 40 countries work together to conduct these assessments. The organization is generally decentralized and, at present, its headquarters are in the Netherlands. In 1995, the TMISS was administered to approximately 500 000 students, in 30 languages, in five grades, and in 45 countries. In Canada, approximately 50 000 students from 800 schools participated in some part of the assessment. Three populations of students were tested. Population 1 included students from grades three and four. Population 2 included students from grades seven and eight. These grades were chosen because they were where 90% of the nine-year old and 13-year-old students could be sampled. Population 3 included students from their last year of secondary school. In Canada, students in Populations 1 and 2 participated either in a written assessment or a performance assessment. Approximately 25 countries participated in the performance assessment portion of the TIMSS. The TIMSS is an international indicators system for science and mathematics. A written assessment of mathematics and physics was administered to students in Population 3; no performance assessment was available for this population. The written assessment consisted of both multiple-choice and extended-response questions. The performance assessment included a series of nine timed practical tasks. Nine students were selected to write in each school. Each student wrote three of the nine tasks in one and one-half hours. At the end of half an hour, the students rotated to their next assigned task. The same tasks were administered to students in Populations 1 and 2, with slightly more sophisticated questions for students of Population 2. A series of in-depth questionnaires were administered to teachers and participating students. These questionnaires allow the researchers to determine the context for teaching and learning in that particular education system. A curriculum analysis of all participating schools is underway. The curriculum analysis and the questionnaires serve as additional indicators for monitoring science education. The report on the indicators studied for Population 2 is expected to be published in the winter of 1996 and in 1997 for Populations 1 and 3. The TIMSS report is expected to include results on student

achievement, a curriculum analysis, student attitudes, teacher backgrounds, classroom organization and instructional practices, and school and school system comparisons for all the participating countries. The context that the students are situated in are factored into their performance on the assessment, therefore rankings of all schools will not be possible.

The International Assessment of Educational Progress

The International Assessment of Educational Progress is a large-scale international comparative educational research project. It is operated under the direction of the Education Testing Service in Princeton, New Jersey, USA. The Education Testing Service is an educational research organization which employs 2 000 highly skilled researchers and statisticians. Generally, five to six industrialized countries cooperate in the International Assessment of Educational Progress assessments. Students aged nine and 13 have been tested to determine their achievement in mathematics and science and geography. In 1991, the second International Assessment of Educational Progress assessment focused on mathematics and science. The traditional International Assessment of Educational Progress assessments are pencil and paper tests. The Education Testing Service has deemed these instruments to be inadequate for assessing the process skills such as observing, measuring, manipulating, and recording, analyzing, and interpreting data that are an integral part of the disciplines of mathematics and science. As a result, the Education Testing Service decided to include performance testing in mathematics and science of their 13 year-old cohort. For the performance testing in science, several short tasks were field tested in May of 1990. Based on this, eight tasks assessing mostly physical science and nature of science concepts and skills were tested on a sample of students from four countries and five Canadian provinces. The Education Testing Service (1992) found that although performance testing can be used reliably, it is three or four times more expensive than pencil and paper tests. Students and teachers reacted positively to this new style of testing, but the standardization of equipment and materials was difficult across the participating countries. The

performance of the students in each of the nine participating provinces or countries was reported per task as a value from 0 - 100%, for example: ". . .78 to 93% of the students categorized . . . correctly" (Semple, 1992, p. 36).

Experience, findings, and reporting strategies, such as these, were important to the SAIP science development team as the framework and the criteria were developed for the SAIP science assessment instruments.

The Assessment of Achievement Programme

Scotland, like other industrialized countries of the world, has developed a national assessment program, the Assessment of Achievement Programme. Unlike many of the other national assessment programs, Scotland's was one of the first to assess the scientific enquiry skills in practical tasks. In 1993, the third Assessment of Achievement Programme survey of student achievement of science was carried out using written and practical assessment instruments. Also, during that year a feasibility study was carried out to determine if talking and writing skills[6] could be assessed in the context of science tasks. The Assessment of Achievement Programme survey was administered to three groups, primary 4, primary 7 and secondary 2. Each student performed a series of practical tasks which assessed skills such as planning, investigating, measuring, collecting evidence, observing, following procedures, processing data, making inferences, applying knowledge, and communicating. Using a criterion-based rating scale the level of the students' response was rated. A student displaying the highest the level of achievement demonstrated sophisticated skills and knowledge. Level of student achievement is reported task-by-task.

The SAIP science assessment, similarly, incorporates the assessment of science inquiry skills and Canadian students also are assessed according to the levels of their

[6]This is an interesting addition; an integrated testing program. This may be the next direction taken in national assessments. Communication is an important consideration in assessing whether a student 'knows' science or any other discipline for that matter.

achievement. Assessment programs, such as the Assessment of Achievement Programme's performance assessment in science served as support for the introduction of a practical component into the science assessment.

Assessment Strategies in Science

Shavelson (1990) observed that comparison-results from international achievement tests have the potential to focus attention and influence the behaviour of policy-makers, teachers, students, and parents. Published results of achievement tests and diploma examinations have the same effect in Alberta. Shavelson (1990) continues by commenting that assessment instruments such as these are an efficient, cost-effective means of providing information on a legitimate, but limited aspect of science learnings. He suggests, as does Wiggins (1989 & 1993), that if test instruments are so powerful, designing a new indicator instrument would effect an immediate response in teaching and learning of science. An instrument that assesses science concepts, science skills, and problem solving represents a more authentic assessment of the teachings and learnings of science (Wiggins, 1989 & 1993) and hence is a more authentic indicator of the student achievement in science. Shavelson (1990) takes into consideration the cost of large-scale practical assessment projects so he calls for research. Indicator instruments should be reformed such that they assess more of what is representative than what is desirable to "know" and to be able to do in the discipline of science. He suggests varying symbolic questions with qualitative expected responses and word questions with quantitative expected responses, reducing the proportion of selected-response questions on examinations, and developing surrogate practical tests, using a computer or a pencil and paper version of the practical task. Shavelson (1990) believes it is possible to develop creative indicator instruments that will exert a positive effect on science education.

Summary

The major education indicator programs and instruments examined in this chapter capture what others view indicators, indicator systems and indicator instruments to be. Indicators are derived statistics that are compared to some reference point. Indicator systems are made of a set of interrelated indicators. Indicators and indicator systems can be developed for an educational system or for even a particular discipline. For example, two indicators for the performance of students studying science are their results on a written and a practical component. Indicator and indicators systems provide information about the performance of education systems. The specific effect of an indicator system or even an individual indicator can be felt when the results of achievement or international tests are published. That the use indicators and indicator systems has such a profound effect on teaching and learning, serves as a reminder of the importance of the type and form that the indicator instruments take.

Chapter III

RESEARCH DESIGN AND METHODS

Introduction

This research study was designed as an historical study which explored the establishment of the SAIP and the development and the implementation of its science assessment instruments. It traced the development of the SAIP from 1988 until 1996 through documentary analysis, participant observer, and semi-structured interviews. A part of the study is an analysis of the mechanisms for the development and implementation of the science assessment instruments. These data formed the basis for the identification of emerging themes and implications of the establishment and development of a national indicators system and of an indicators instrument.

Overall Research Design

This study is based on constructivist premises that people provide potentially unique understandings of activities undertaken in common. Personal experiences, formal education, social status, and gender may all influence these understandings, which are the reality of the people involved. Further, these understandings may be reframed or reconstructed in discussion with the researcher and will be influenced by the social position of the researcher herself and her contributions to the conversation. As a participant observer, the researcher has a history with each interviewee and also her own construction of events. During the conduct of the study, the researcher tried to adopt an open stance to alternative constructions of events rather than to seek one coherent story. This involved both bracketing her prior conclusions and probing for a detailed understanding of others' stories.

In an historical study, even of the recent past, the major data sources are artifacts and memories. The research project unfolds through a series of iterations. First, a

preliminary documentary analysis provides preliminary factual data and publicly acceptable resolutions of agenda items. Interviews with people involved with the SAIP, enrich and challenge this initial framework sending the researcher back for further reading of documents and in search of other informants who could provide additional data, whether contradictory or confirmatory. As a participant observer, the researcher is more conscious of observing and documenting incidents and routines and testing these interpretations in discussion with other participants. Spradley (1979) refers to this process of question, interview, analysis, and further questions as the ethnographic turn. Specifics of the research design including data sources, data collection, and data analysis follow.

Data Sources

The artifacts of the SAIP include the audio tapes, minutes, documents of official and subcommittee meetings at the local, provincial, and national levels, the published and unpublished reports, and the memories of the participants of the key committees.

Meeting Minutes and Audio tapes

Through the government libraries and the archival storage system, the minutes and briefing notes for the ACDME and the CMEC meetings from 1987 through 1996 were located. A preliminary reading was made of these. However, careful notes were made from the sections of the minutes which dealt with the SAIP and the 'education indicators program' as it was first known. The SAIP was on the agenda of every meeting of Deputy Ministers and Ministers between 1988 until 1996. Recommendations from working committees and groups were made to the deputy ministers and upon their approval, final approval was sought from the Ministers' of Education. Once these documents were researched, a final search of the CMEC library in Toronto was made in July 1996 to ensure reliability and accuracy.

The proceedings of two meetings were available on audio tape. The teleconference of the Science Development Team, held on July 7, 1994 and three hours of the in-service for the administration of the practical component of the science assessment, held on February 21 and 22, 1995, were audio taped. These tapes were listened to and relevant information recorded.

CMEC Publications

The CMEC Annual Reports from 1974-75 until 1992-93 and the Biennial Report of 1993-1995 were read. No information about an indicators program was published in the issues prior to and including 1987-88 Annual Report. From 1988-89 until and including the Biennial Report, every report included a summary of SAIP activities.

Issues of Liaison dated as early as 1986, another official CMEC publication, were obtained. It was published twice a year until its publication was suspended in the fall of 1991; it was reintroduced in the spring of 1995. The January 1986 issue reported a joint meeting of CMEC officials, CERI experts and government of Alberta personnel held in Edmonton, Alberta. The purpose of this meeting was to study exemplary practices for the transition of the handicapped from school to work. Although not directly on the topic of indicators, connections were made between statistics gathered on the outcomes aspects of education such as transition from school to work and the SAIP as it was initially planned. Educational indicators were first mentioned in this publication in the December 1988 issue, as a one line mention. In June of 1989, now called the CMEC School Achievement Indicators Project, 20% of the document was devoted to this new topic. Thereafter, a section of every issue dealt with indicators.

The public reports of the mathematics and the reading and writing assessments were read, as was the CMEC Report on Education in Canada, released in November 1995. A series of unpublished documents prepared specifically for the establishment of the SAIP

and the science assessment were also used as reference materials and are referred to in Chapter 4 of this study.

Personal notes

The personal notes of one of the key informants who served on the Alberta Project Team[7] were made available as reference materials. The researcher was provided with six binders of hand-written notes, personal records of meetings and teleconferences, and unpublished documents. These were an invaluable source of information for helping to trace the history of the establishment of the SAIP from 1989 until 1990. They enhanced the briefing notes and official minutes by providing a richer, more in-depth perspective.

As participant-observer, the researcher followed up each meeting with carefully recorded field notes of the proceedings. These notes were kept with her SAIP files that she established in chronological order. Between meetings, she kept files of all pertinent documents produced or used in the development of the science assessment.

Documents

An attempt was made to review all available SAIP-related documents in various government libraries in Edmonton and at CMEC headquarters in Toronto. Then, pertinent documents were summarized and analyzed. This helped to determine the chronological history of the SAIP and the science assessment and further, to answer some of the research questions. This document analysis also helped the researcher to begin a list of potential key informants.

the text continues

[7]The Alberta Project Team was set up by the Director of Student Evaluation and Records Branch in 1989 to prepare documents for the consideration of the Technical Committee on SAIP. Six members of Alberta Education sat on this team. It was disbanded in June 1990; the same time as the Technical Committee on SAIP was.

Participants

The initial interviewee was identified as a key informant because she was a writer of significant documents which helped establish the SAIP. She was identified through the first document analysis. Subsequent informants were identified through document identification, the 'snowball' interview technique, and personal knowledge. Nine key informants were formally interviewed and their conversations recorded. These interviews were followed up with informal conversations which helped clarify areas of uncertainty. The key informants, then, are a purposive sample from those involved in the SAIP. By participating in the interview implied consent was obtained. All interviewees were assured of confidentiality, anonymity, and the right to opt out of the study at any time without reprisal. Participants were made aware that they would be referred to in this study, according to the position they held and the jurisdiction or organization they represented. A brief description of each of the key informants follows.

Deputy Minister of Education of Alberta

The Deputy Minister of Education of Alberta was a member of the Advisory Committee of Deputy Ministers of Education from 1982 until 1995. As a member of the ACDME, he served on several action committees as well as serving as the chair of the ACDME in 1984-85. As the chair of the ACDME Subcommittee on Priorities and Programs - Elementary and Secondary Education (ESE), he presented the idea of an indicators project to the deputies and to the Ministers in 1987-88 and again in 1988-89. His main responsibility as it related to the SAIP was to ensure that a meaningful indicators program be developed. He did this by mobilizing his staff at Alberta Education to write the proposals, goals, objectives, and action plans for an indicators program. He was involved with the SAIP from its inception in 1988 until his retirement in 1995. At the Deputy Ministers' and Council of Ministers' meetings, his Assistant Deputy Minister played an active role in explaining the rationale of SAIP to the parade of new Deputies and Ministers

of Education over the years. Hence, Alberta provided the continuity much needed to ensure a project such as this one survive over time.

Chair of the Technical Committee for the CMEC SAIP

The Director of the Student Evaluation Branch of Alberta Education was the chair of the Technical Committee for the CMEC's SAIP from 1988 - 1990. This committee was composed of at least one representative from each jurisdiction. It was created to establish the education indicators program. As the committee chair, he worked with the representatives to arrive at consensus to determine the exact form that the education indicators would take. The decisions were presented as recommendations to the ACDME. Prior to chairing the Technical Committee, he played a key role in developing the original concept of the indicators program. He was also the chair of the Ad Hoc Test Development Group in 1990. By meeting the Directors of Evaluation from the jurisdictions, he and others recognized the need for such a program and the work began, in a formal sense, at Alberta Education in late 1987 or early 1988. Besides chairing these key committees, he was responsible for the Alberta Education staff members who played a central role in the development, administration, and standard setting of the SAIP Math, Reading and Writing, and Science Assessments.

Alberta Project Team, member

The Assistant to the Director of Evaluation of the Student Evaluation Branch of Alberta Education was a member of the Alberta Project Team from 1989 to 1990. This six member Alberta Project Team was established by the Director of the Student Evaluation Branch of Alberta Education. This team wrote position and methodology papers as proposals to the Technical Committee and thereby helped shape the education indicators program. Along with other responsibilities, the Assistant to the Director of the Student Evaluation Branch of Alberta Education wrote a position paper on literacy that helped

establish the rationale for the Reading and Writing Assessment. After her involvement with this Project Team, she occasionally reviewed materials that formed the basis of the Reading and Writing Assessment. She drafted news releases and briefing notes for Alberta's Deputy Minister and Minister of Education for ACDME and CMEC meetings, respectively. She responded to questions in the department and helped review draft materials on the SAIP for department publication.

Technical Committee member and Science Contact

This interviewee joined the Technical Committee on SAIP in the fall of 1989, as the Ontario representative and remained a member until it was disbanded in 1990. Prior to that time, he was employed by the Ministry of Education in British Columbia but did not participate in the national meetings. He was a member (representing Ontario) of the Ad Hoc Test Development Group in 1990 along with representatives from Alberta, British Columbia, Québec and Prince Edward Island. The mandate for this group was to determine whether the assessment instrument should be purchased or developed. In 1991, when Ontario joined the Consortium for Developing the SAIP Assessment Instruments, he was named as this province's representative, and concurrently he was the provincial coordinator for the SAIP Assessment Instrument Development. For Ontario, he was Project Director of the administration of the SAIP Assessments and as part of the Consortium, he served as the national coordinator for the administration, marking, and data analysis of the Math and Reading and Writing Assessments. He provided information to the Report Development Group about the Math and Reading and Writing Assessments. In December of 1993 he was named as the Ontario provincial coordinator (Science Contact) for the SAIP Science Assessment. He served in these various capacities until he left the Ontario Ministry of Education in the fall of 1994.

Report Development Group member and Science Contact

This interviewee was responsible for the administration of the SAIP Science Assessment for Ontario, from 1994 until 1996. He made recommendations about the administration, scoring, and report writing of the assessment. As well, he was the Ontario SAIP provincial coordinator (Science Contact) for the science assessment and was the interim provincial coordinator (Math Contact) for the 1997 Math Assessment. He was also the chair of the Report Development Group for the Reading and Writing Assessment in 1995. Prior to September 1994, as Director of the Assessment and Evaluation Branch of the Saskatchewan Department of Education, he acted as the provincial coordinator for the Science Assessment (Science Contact). In the formative years of the SAIP, he was an advisor to the BC representative to the Technical Committee in 1988-89.

Report Development Group, report writer

This interviewee was one of two report writers of the public and technical reports for the math, reading and writing, and for the science assessment. As report writer, he met with the Report Development Group. This group set the framework of the public and technical report and made recommendations to the SAIP Policy Advisory Committee. He has been involved with the SAIP report writing from 1991 until 1996. Prior to this involvement with the SAIP, he was the Director - ESE as part of the CMEC secretariat, for seven years ending in 1986-87. Because of this experience with the CMEC, he was hired as the political analyst for the public report and the person who would write the French adaptation. During the writing of the reading and writing, he became the principal writer, and is the principal writer of the science public and technical reports. During the years he spent with the CMEC, the idea of an indicators program had been discussed by the Subcommittee of Priorities and Programs - ESE as the possibility of setting national standards for student achievement in education.

National Coordinator of the SAIP

This interviewee was the National Coordinator of the CMEC School Achievement Indicators Program from March of 1994, when the position first came into existence, to the present. Prior to March 1994, the duties associated with the SAIP were carried out by the Coordinator - ESE within the CMEC secretariat in 1988-92, and this position from 1992-94 was Coordinator, CMEC secretariat, with responsibility for the SAIP. The National Coordinator was responsible for coordinating activities related to planning and implementing the SAIP. She ensured smooth communications between the CMEC secretariat, provincial and territorial representative, the Administration Management Team, the Report Development Group, and the assessment Development Team. She managed the SAIP budget, assisted with managing the Development Team budget, reported in-kind and indirect expenditures by jurisdictions, and maintained liaison with other CMEC coordinators in their work with the Report on Education in Canada and the Pan-Canadian Education Indicators Program. She attended one day of almost every Science Development Team meeting.

Team Leader, Mathematics and Science Development Teams

This interviewee was the leader of both the mathematics and the science assessment Development Teams. He worked on the mathematics assessment from the fall of 1990 until the completion of the instruments in December 1992. His work continued again with the mathematics assessment with the following responsibilities: preparation of the technical information booklet by providing the field test statistics and the assessment instrument structure, in March 1993; the establishment of the short-answer and the problem-solving scoring guides; scoring of the short-answer component, in June 1993; involvement in one of the three regional scoring sessions of the problem-solving component in July 1993; and was present in Quebec for the release of the public report on the mathematics assessment.

From 1990 until 1993, he represented Alberta as the provincial coordinator or Provincial Contact as it was known, then for the SAIP mathematics assessment.

His involvement with the science assessment began in August 1993 when interprovincial discussions with key science educators were held. He was instrumental in writing, on behalf of Alberta, the four part proposal that laid the foundations for the SAIP science assessment. The writing and presentation of this proposal to the ACDME and the CMEC in the fall of 1993 and its subsequent approval was the result of his work. The responsibilities of the Development Team Leaders for both the mathematics and science assessment had some similarities. These were to coordinate the activities of the Development Team in the development of the assessment design, framework and criteria, items, field tests and final instruments, and communication documents. As the leader of the Development Team, he prepared the budget, timeline, agenda, chaired meetings and moved the project along. His responsibilities with the Development Team for science assessment began in 1993 and concluded in August 1996. These responsibilities included the supervision of the administration of the science assessment. When the second administration of the mathematics assessment began in November 1995, he was a member of the steering group that initiated the process and established the formal Memorandum of Understanding. He will likely be involved in a similar consultative role when the second administration of the science assessment is considered. As the standard-setting and the report writing gets underway in the fall of 1996, he likely will be involved in an informative and consultative role.

Science Specialist - Practical Tasks

This interviewee was one of six Science Specialist-members of the Development Team for the science assessment from March 1994 to January 1996. He was one of two representatives from Saskatchewan. In this case, this Science Specialist coordinated the production of stations for the practical component of the science assessment. He was

responsible for reviewing, redesigning and formatting the written materials for the practical tasks. He was also responsible for acquiring materials and supplies needed to assemble the stations and their subsequent packaging and shipping to all provinces and territories. He wrote the "Handbook for Practical Task Coordinators" to assist those administering this component of the SAIP science assessment. After leaving the Development Team to be the Assistant Director of Assessment and Evaluation of the Saskatchewan Department of Education, he continued his involvement with the SAIP by acting as the Provincial Coordinator for the Saskatchewan oversample of 2 000 students in the practical component of the science assessment and he acted as the provincial coordinator for Saskatchewan for the SAIP Science Assessment (Science Contact) for the remainder of this assessment cycle.

Science Specialist and Participant Observer

As the researcher of this study was a member of the SAIP science assessment Development Team as a Science Specialist, she also gathered data by observation. Her formal involvement spanned from June 25, 1994 to August 2, 1996. As part of the Development Team, her responsibilities included the development of the assessment design framework and criteria, preparation of field tests and final assessment instruments and writing the documents associated with the scoring and administration of the assessment. Along with a Science Specialist from Ontario, she coordinated the scoring of the written component of the field tests in June, 1995.

A Pilot Study

A pilot study was carried out in the spring of 1995 to confirm and refine the methods planned. It served as a feasibility study. A few published CMEC documents were reviewed, a list of possible key informants was drawn up, and one key informant was interviewed. The interview was then transcribed and analyzed. These data were subsequently used in the main study.

Data Collection Procedures

The data were obtained from three major sources: documents, interviews, and by observation. Although not linear as depicted, the clearest way to explain the circuitous data gathering procedure used in this study is as a list. However, the description of interviewing procedures and the data analysis section of this study explain the wandering, repetitive, data gathering procedure actually used by the researcher (Spradley, 1979).

The following data collection procedures were used, at some time, during the data gathering stage of this study.

1. The researcher used GATE Search and ERIC Search to find pertinent information and location of resources on indicators, indicator systems, science assessments and the SAIP.
2. The researcher used the telephone and e-mail systems to obtain annotated listings of the documents available.
3. The researcher read all official meeting minutes of the ministers and deputy ministers of education, from 1988 until 1995, (the ones in 1996 were not available) and summarized them.
4. The researcher read available briefing notes, unofficial meeting minutes, and listened to the available audio tapes and summarized them.
5. The researcher reviewed the information and chose the relevant documents, obtained copies as required, then read and summarized them.
6. The researcher compiled a list of key informants who served as participants of the study.
7. The researcher contacted the key informants to obtain permission and to arrange interview times by telephone or in person and followed up with a letter.
8. The researcher interviewed the key informants using the interview guide, focusing on their particular area of involvement in the SAIP.

9. The interviews were transcribed and then mailed to the informants for their changes, additions and/or deletions.
10. The researcher analyzed the data, by finding basic historical data and by picking out categories and then emerging themes.
11. The researcher discussed findings with select key informants to increase her understanding of the complexities of the interactions.
12. The researcher made and recorded observations of the proceedings of all Science Development Team meetings which she attended; the proceedings were summarized as part of the data analysis. The researcher kept chronological files of the proceedings of the SAIP science assessment for the extent of her involvement.
13. The researcher discussed the summaries and draft versions of the findings chapter with another team member, the team leader, and the National Coordinator, as was necessary and at different times.
14. The researcher asked that a draft version of chapters one through four be read by the statistician for the SAIP science assessment and by an analytic resource officer for Alberta Education.
15. A accurate record of these procedural steps incurred in this study was kept. Should the need arise, an audit of the research is possible.

Description of the interview process

Once the researcher decided who the participants of the study were, most of them were approached in person and invited to be participants by being involved in an interview. The interview time and location were discussed and decided, mutually. Two of the participants were contacted by telephone first. The study was explained to the participants verbally and a brief description was provided in writing as part of the follow-up letter that each participant received (See Appendix A). One key informant was interviewed in a hotel

room in Toronto and another was interviewed in a meeting room at the University of Alberta, however, most of the key informants were interviewed in their own offices or at their place of work. A semi-structured interview guide (See Appendix B1) with several open-ended questions was used. The problem statement and the research questions of this study guided the structure of the interview guide. However, the emphasis of the interview depended upon the role the interviewee played in the SAIP or science assessment. All formal interviews were audio taped and then transcribed. The transcriptions were mailed to each participant for verification; a thank-you paragraph was included (See Appendix C). A follow-up informal interview (See Appendix B2) was held with informants to determine, from their perspective, what they believed their roles and responsibilities were, with respect to the SAIP. Most of the follow-up interviews took place over the telephone.

The first interview of the nine participants was held in February 1995 and the final interview was held in June 1996, when the list of participants was, finally, finalized. Each interview lasted approximately one hour, however, one interview was only half an hour long, another was longer than the one hour tape that the researcher had brought along (the remainder of the interview was recorded by the interviewer as notes), and one interview was almost two hours long. With the exception of one participant, the researcher knew, worked with, or was a student of, all the participants of the study.

Data Analysis

A systematic analysis of historical documents, transcribed interviews and field notes was done to identify their messages. The analysis was ongoing and recursive (Spradley, 1979). Layers of work form the findings of the study. This content analysis was guided by the criteria established by the problem statement and research questions of this study.

The data from publications such as the CMEC Annual Reports and the Liaison form the chronological "backbone" of the study. This analysis gave the outline of the

development of the SAIP over time. Data from meetings, either minutes, audio tapes, or field notes, provided factual information about how the establishment of the SAIP progressed and about how the science assessment instruments were developed.

Categorical analysis of the written messages of the taped interviews was coded out using the open coding procedure described by Strauss (1987) in Berg (1995, pp. 186-188). The main categories that emerged were the units of analysis. Then inductive reasoning was used to derive meaning from these categories and to identify themes. Hence open coding, categorization, thematic content analysis, and inductive reasoning outline the data analysis procedure used on the transcribed interviews. However, the data gathered from the interviews were integrated into and only slightly extended the data gathered from the document analysis.

The study findings chapter was written as a series of three printed drafts, completed June 30, July 30, and August 31, 1996, respectively and numerous drafts stored on computer disks. Each draft reflected additional layers of data gathering, analysis, and confirmation.

Delimitations

This study was delimited to (a) major aspects of the SAIP, (b) the SAIP science assessment, and (c) the time frame from 1988 to August 1996. Of all the contributors to the SAIP, the researcher chose nine key participants for interviews. The major aspects of the SAIP investigated were its genesis, rationale, development, leadership and implementation.

The science assessment was the one example of the implementation of the SAIP examined in detail. The mathematics and reading and writing assessments were researched only as they related to the establishment of the SAIP. While essentially the study was concerned with the years 1988 until 1996, preceding events were researched to determine the genesis of the SAIP.

Limitations

The design of this study required that the writer examine documents, interview key informants, and participate in observations. A potential limitation of the document analysis was the problem of doing an exhaustive document search, since pertinent documents were located at the various Ministries of Education, at the CMEC head office in Toronto, and at Statistics Canada in Ottawa. Interprovincial loaning took place with the hope that as many of the documents relating to the SAIP as was possible were read and analyzed. However, no participant referred to any documents which the researcher was unable to locate and read. The interviewing process involved the problem of distance, therefore the interviews of the key informants from Toronto and Regina were held concurrent to Science Development Team meetings held in Toronto. Follow-up discussions were held either by long distance telephone calls, e-mail, or in-person at subsequent meetings. A potential limitation involved the memories of the interviewees and the skills of the interviewer. While the researcher found that the participants seemed to enjoy talking about the SAIP, she was conscious that as a participant and relatively recent member of the SAIP process, she may have missed references to processes or issues, or not have been informed of issues which could reflect poorly on the SAIP. There were two such occasions when the researcher felt more information could have been forthcoming; one, the early Ontario non-involvement with the SAIP and the other, the Saskatchewan non-involvement in the mathematics and reading and writing assessments. Although interesting, these issues were not actively pursued by the researcher, since they appeared to resolve themselves over time.

Acronyms

Because this study involves the discussion of many known projects and programs conducted by known organizations, the use of acronyms is necessary. These acronyms have been compiled and are presented in Appendix D.

Trustworthiness of the Findings

The researcher used three methods to explore the same problem statement. The combination of document analysis, the series of interviews, and being a participant observer served as a form of triangulation and strengthened the research findings. The systematic and extensive review of public documents make the research findings dependable and credible. Specifically, upon near completion of the study, the researcher traveled from Edmonton to the office of the CMEC secretariat in Toronto to read the official minutes of the meetings of the ministers and deputy ministers of education. The readings and notes were compared to the findings and no significant inconsistencies were found. However, compared to official minutes, the study seemed to provide more information. To test the credibility of this apparent "extra information" one-on-one discussions were held with pertinent key informants and other members of the Science Development Team. Specifically, questions about the CMEC approval process were asked of the National Coordinator, the Science Development Team leader, and the Science Specialist from Nouveau-Brunswick. As a similar story emerged, the information was retained. Also, two people involved in different aspects of the SAIP read the findings of this study. One, the statistician, was actively involved with the SAIP in its entire duration from its establishment to the analysis of the statistics obtained from the mathematics, reading and writing, and science assessments. In an informal interview, in a page by page discussion, he provided clarifications and suggestions for inclusion of information he deemed missing. His overall comment was that the findings read "just like a movie." The other, an analytic resource officer from Alberta Education, was involved as a member of the Alberta Project Team from 1988 until 1990; the team which helped establish the SAIP. His comments were relatively minor and were incorporated into the final draft of this study.
Collectively, these procedures increase the confidence, credibility, dependability, and trustworthiness of the findings.

Ethical Considerations

An application for permission to conduct this study was sought from an ethics review committee of the Department of Educational Policy Studies of the University of Alberta. An affirmative reply was received. Letters, one to the province of Alberta and another to the CMEC secretariat, requesting permission to use government documents and interview government personnel were sent. Both, the province of Alberta and the CMEC secretariat replied affirmatively (See Appendices E1 and E2). Permission was also requested of the key informants. These informants were assured of confidentiality, anonymity, and the freedom to opt out of the study at any time, in writing (See Appendix A). Although much of the information gathered was from public documents, personal notes obtained from one key informant were treated with confidentiality and then returned. The researcher was a team member of the Science Development Team, however, as much as possible work as a member of the Science Development Team was kept separate from work as a researcher. All interviews were held on the researcher's study time and not on her work time. Also, members of the Science Development Team were cognizant that the researcher was gathering information on the developmental process of the SAIP science assessment and gave their permission for her to use information gathered during her work.

Chapter IV

FINDINGS OF THE STUDY

Introduction

This chapter presents, in chronological sequence, the genesis of, the context for, and the establishment of a national indicators program in Canada. This program is the SAIP - School Achievement Indicators Program. It is one of the major programs operated under the direction of the CMEC - the Council of Ministers of Education, Canada. Therefore this chapter begins with a brief description of the CMEC, its organization, its operating structure, its objectives over time, and its official and unofficial functions in education in Canada.

This chapter also traces the development and eventual administration of the SAIP science assessment from its conception in 1993 to the report writing stage in 1996. How the framework, the criteria, the instruments, and the supporting documents were developed is explained, in a chronological sequence. Additionally, the statistical instrument and sampling designs, the administration and scoring procedures, and the plans for the standard setting processes are explained.

Finally, the chapter identifies some themes common to the establishment of the SAIP and to the development of the SAIP science assessment.

A chronological summary of the events starting from the genesis of the SAIP until the administration of the science assessment can be found in the next chapter.

THE DEVELOPMENT OF THE SAIP

The CMEC

CMEC is an organization established in 1967. Although, education in Canada is the responsibility of the provinces and territories, CMEC is body that facilitates harmony, cooperation, collaboration, and information sharing among these jurisdictions at a national

level while allowing them to maintain their independence. It provides other nations a single body with which to communicate on matters of Canadian education, it represents education in Canada in international settings, and it is the official channel for decision-making on topics concerning Canadian educational policy. The Council also serves to support the cultural and linguistic duality of Canada (CMEC, 1988b & 1994a).

The members of the CMEC include all Ministers of Education representing the provinces and territories[8] of Canada. The ministers are responsible for elementary, secondary, and or post secondary education in Canada. As of 1996 there were 12 Ministers of Education and four Ministers of Advanced Education for a total of 16 members of the CMEC. Members may be represented by their deputy minister or other named designate, at the meetings of the CMEC. Members, or their designate, can be accompanied by ministry advisors as members deem necessary (CMEC, 1994a & 1996).

The members of the Council meet at least two times a year, in a chosen jurisdiction, to discuss matters of mutual concern. Generally the meetings are held in February and September of each year, with one meeting in Toronto and the other by invitation in another jurisdic[illegible]. The ministers are seated from west to east and the flags of the jurisdictions are set up [illegible]der that the jurisdiction entered confederation. As of the 1994 constitutional revision, the chair is elected biennially[9] and the vice-chair, annually, from the members of the Council. The meetings are conducted in both official languages, English and French, with simultaneous translation provided for those who require it (CMEC, 1994a).

An Executive Committee, which elects its own chair from among its members, consists of at least five identified (by election or by designation) members of the Council. The Committee includes one member from each of the five regions of Canada: British Columbia and Yukon, the prairies (Alberta, Saskatchewan, Manitoba, and the Northwest Territories), Ontario, Québec, and the Maritimes (New Brunswick, Nova Scotia, Prince

[8]At the 65th meeting of the CMEC held in Toronto on February 28, 1994 the territories gained membership to the CMEC. Prior to that time they held observer status.
[9]Prior to 1994, the chair of the CMEC was elected annually.

Edward Island, and Newfoundland). It exists to serve as a steering committee, finance committee and or an action committee in cases of urgency, on behalf of the Council. Generally, meetings of the Executive Council are held one day prior to the meetings of the Council (CMEC, 1994a).

Also outlined in the Agreed Memorandum on a Council of Ministers of Education, Canada (1994), is the Advisory Committee of the Deputy Ministers of Education (ACDME). The membership of the ACDME includes all the deputy ministers of the members of the Council. Like ministers, deputy ministers may be represented by named designates and deputy ministers or their designate can be accompanied by ministry advisors to meetings of the ACDME, as deemed necessary. Generally, the ACDME also meets two times a year. Its meetings are held in January and August; one month before the meetings of the CMEC. The meetings of the ACDME are held at the office of the secretariat of the CMEC in Toronto. The major responsibilities of the ACDME are to review all matters put forward to the Council, makes recommendations to the Council on action items, bring forward issues to allow the Council to formulate policies and proposals, and to prepare the agenda for the next meeting of the CMEC.

Members and representatives of members or their delegates may form sub-committees of the Council and the ACDME to execute the responsibilities of the Council.

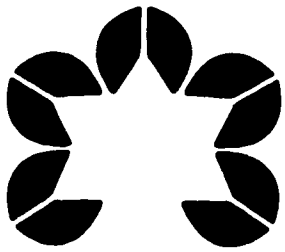
CMEC also includes a secretariat that provides a central office and functions to research, coordinate, administer, distribute materials, and to carry out duties assigned to it by the Council, including assisting the Council and any of its committees. The secretariat is incorporated and is known as THE CORPORATION OF THE COUNCIL OF MINISTERS OF EDUCATION/LA CORPORATION DU CONSEIL DES MINISTRES DE L'ÉDUCATION, CANADA. The Council hires a Director General to act as the chief executive officer, treasurer, and secretary of the Council. The Director General of the CMEC is an ex-officio member of each committee or task force of the CMEC. The

secretariat services are provided by a group of approximately 30 people. It is based in Toronto and operates in both official languages (CMEC, 1994a).

In 1996, the Chair of the CMEC was the Honourable Gordon E. MacInnis, Minister of Education for Prince Edward Island. His Deputy Minister, Mr. Mel Ostridge, was the Chair of the ACDME. The Director General of the CMEC, newly appointed in July 1996, was Dr. Paul Cappon, formerly Vice-President, Academic, Laurentian University in Sudbury, Ontario (CMEC, 1996).

The activities of the CMEC and the operation of the secretariat are mostly financed by the provinces. The income of the CMEC depends upon contributions of the provinces and territories, a federal contribution, and some private contributions. Major projects such as the SAIP have an independent financial arrangement (CMEC, 1994a).

Figure 4-1 The CMEC Logo



Source: CMEC secretariat

Illustrated in Figure 4-1 is the logo of the CMEC. The 10 individual pieces of the logo represent the 10 provinces. The five pairs of pieces represent the five regions of Canada and the centre star of the logo represents the maple leaf and the Centennial star, the symbol of the 100 year celebration and also happens to be the year of the birth of CMEC.

What are CMEC's objectives?

Over the duration of its existence, the CMEC has had various main objectives. In 1967 these were interprovincial consultation, federal-provincial negotiation and provincial-international participation. In 1983 the CMEC underwent a major reorganization. A reevaluation of the Council's role and operating procedures began in 1981-82. In 1982-83

the Council altered its structure, operating procedures and set out new goals and objectives. All Council committees prior to September 1983 were dissolved except those with a joint CMEC-federal mandate. Two Subcommittees of the ACDME were created; the deputy ministers were mandated to take control of the Council's programs and keep them on track. The Council reaffirmed its interest in elementary, secondary and postsecondary education and increased the ministers' direct involvement in its activities. The new goals and objectives included a plan for increased coherence and direction in its work, increased visibility, a voice for education, proactive work, developing a national perspective, addressing national concerns, and cooperating with the various stakeholders of education. These objectives strengthened the actions of the CMEC in the eighties. In 1993, the Council again revisited its goals and objectives. This time they were framed in concrete terms and written up as an interprovincial agreement. In September 1993, all the ministers agreed to the Joint Declaration - Future Directions for The Council of Ministers of Education, Canada. This agreement included one major goal--to adopt a national approach to dealing with common educational problems; a mission statement expressed as the national education agenda; and an action plan that listed a series of Council activities that would serve to meet their goal and mission (CMEC, 1974-1993).

The official function

Education in Canada is the responsibility of the provinces and territories. CMEC is a body that facilitates harmony, cooperation, collaboration, and information sharing among these jurisdictions at a national level while allowing them to maintain their independence. It provides other nations a single body with which to communicate on matters of Canadian education, it represents education in Canada in international settings, and it is the official channel for decision-making on topics concerning Canadian educational policy. The Council also serves to preserve the cultural and linguistic duality of Canada.

Various annual reports from 1974-1975 until the biennial report of 1993-1995 describe the function of the CMEC in one or more of the following ways. The CMEC

- provides for interprovincial cooperation and consultation,
- facilitates the exchange of information among and between education ministers,
- undertakes projects of value to all jurisdictions,
- effects intangible but real interprovincial cooperation,
- provides for federal-provincial negotiation, communication, and cooperation,
- provides for international communication,
- uses a joint decision making process to seek consensus so that all decisions are considered to be unanimous,
- is the official channel for decisions affecting Canada-wide educational policy, and
- is the official channel to deal with federal ministries on matters related to representation at international conferences.

Committee work, task forces, and studies are the primary means for achieving interprovincial, federal-provincial, and international cooperation and consultation.

The unofficial function

Jacques-Ivan Morin, the Chair of the CMEC in 1977-78, described the unofficial function of the Council as an adversary of the federal government: "Had it not been for the Council, the federal government, always inclined to extend its jurisdiction, might have made further inroads into the field of education as it has done in cultural affairs." (CMEC, 1978, p. 1) The jurisdictions had gained strength by organizing themselves to retain education as their responsibility.

In an interview reflecting on the early years of his involvement with the CMEC, Bosetti (1996) commented that

> . . . the Council of Ministers was always responding to what I call the call to action. Like Paul Revere riding through the streets of Boston, we have the ministers

responding to the call: "The feds are coming! The feds are coming!" Immediately the Council would say we need to do something. "It's our affair and we're looking after it thank you very much!" In those days the Council did a marvelous job of keeping the feds at bay, especially at keeping the feds from finding any need to establish a formal, National Department of Education.

A search of the minutes of the CMEC does indeed seem to bear out this description of an aspect of their function. The June 1988 minutes report that the Honourable L. Bouchard then Minister of the Secretary of State presented the proposal of a federally funded Canada Scholarship Program. To determine the candidates the Secretary of State suggested the implementation of a national test. In September 1988, a new federal department, the Ministry of State for Science and Technology and a new minister made the same proposal to the Council. The ministers of education argued that a national examination would be a federal intrusion into provincial matters, that it may have a steering effect on curriculum and that adequate methods to choose comparable candidates already existed. Again the federal minister left empty-handed. In February 1989, another new federal minister with the same speech met with the same response. In September 1989, the federal ministry sent a representative to make the same appeal. This time the federal government was invited to try the scholarship program, for one year, without the examination. The idea of a national test came up yet again, at the February 1990 meeting of the CMEC. However, at the September 1990 meeting, the Canada Scholarship Program received full approval of the CMEC without a national test. The CMEC had been "marvelously" effective; its students reaped the benefit of federal funds and the provinces were able to avoid having the federal government design a qualifying national examination.

CMEC's activities

Notwithstanding its role in keeping 'the feds' at bay, the CMEC has been and is involved in many interprovincial, national, and international programs. One of the

programs that has been the most significant in the history of the CMEC is the SAIP. The 1990-91 CMEC Annual Report notes on page six: "The School Achievement Indicators Program, which is developing indicators of student achievement and involvement in Canada's school system, is the most demanding project undertaken by Council in its 25-year history." A description and explanation of the genesis of, the context for, and the establishment of the SAIP follows.

The Genesis of the SAIP

From as early as 1985, the minutes of the CMEC meetings report regular discussions of standards in education. At the 47th CMEC meeting held in Québec City on September 16 and 17, 1985, Alberta announced its creation of a Council on Alberta Teaching Standards. At the 49th CMEC meeting in Winnipeg, September 15 and 16, 1986, CMEC mandated that the Council's secretariat undertake a survey that would describe the scope of student evaluation as a requirement of provincial policy and describe the rationale for implementing and maintaining student evaluation programs. At the next CMEC meeting held in Toronto (February 2 and 3, 1987) student evaluation was on the agenda again. The minutes of the 32nd (1982) to the 46th (1988) meetings of the ACDME reported the discussion of standards in education, student assessment, and education systems evaluation. In 1983 and 1985, high school equivalency testing, evaluation as a priority of the ACDME Subcommittee on Priorities and Programs - ESE[10], and finally key indicators of system behavior, especially student evaluation, were discussed but resulted in little action. In 1987, the ACDME Subcommittee on Priorities and Programs - ESE discussed the concept of education indicators for secondary schools, national standards and a national achievement testing system. The minutes of the 42th meeting of the ACDME (January 13, 1987) reported that the already approved 1987-88 activity on systems evaluation was

[10]This subcommittee and another for Postsecondary Education were established by the ACDME at the September 25, 1983 meeting.

augmented to include a study of the need for national standards and for a country-wide achievement testing program. While the notion of a national testing system did not meet with approval by the CMEC, a feasibility study to help the CMEC determine the need for a national achievement testing system and for national standards was approved. In December 1987, the CMEC decided to put the study under the direction of the ACDME Subcommittee on Priorities and Programs - ESE.

In April 1988, the secretariat of the CMEC conducted a survey to determine what provincially sponsored programs the jurisdictions already had in place to assess student achievement and evaluate education systems. The initial survey had been expanded to include questions about programs that examined individual student performance, group performance, and system evaluation. The research study found that 10 jurisdictions, with the exception of Manitoba and Ontario had external student testing programs; nine jurisdictions with the exception of Saskatchewan, the Northwest Territories, and Prince Edward Island, had programs that surveyed overall student performance by sampling student achievement; and seven jurisdictions had formal programs to evaluate the effectiveness of educational programs by using a variety of indicators (achievement, instructor qualification and evaluation, resources available, public satisfaction, etc.) (CMEC, 1988).

A critical analysis was done to determine if there were national needs that could be met with a CMEC sponsored national action. One major finding was that, although sound provincial or territorial sponsored student achievement programs existed, they related to local standards and had limited capacity to relate local results to broader standards. The report noted that the programs were adequate but to be able to improve or fine tune the performance of the education system, educational performance standards needed to be identified. These standards would be national, not regional or local, and would serve to prepare Canadians for the emerging global information age. National educational standards were seen to be more useful if developed in tandem with an evaluation system which would

enable the level of students' performance to be measured against them. With this type of evaluative strategy occurring with continuity, judgments could be made about the adequacy of the Canadian education system and it would ultimately provide a basis for establishing a mechanism for public accountability. Local programs and standards were not seen to increase Canada's competitiveness in the global economy that was emerging. These were presented as arguments for the establishment of a national educational evaluation system. The report also argued that since the Provinces are responsible for education, a national endeavour such as this should be coordinated by the CMEC. Further the national standards program would allow for the calibration of local programs. It was suggested that the proposed program include a Canada-wide testing program in language, mathematics and science. These were considered to be the best indicators of overall system performance, most important for further education, and most significant, relative to national priorities.

The report contained a further series of recommendations. First the testing program should be criterion-referenced in order to determine the level of student achievement in relation to specific criteria. Then, the language test should be based on criteria that would outline the level of literacy functionally required for everyday Canadian life. The science and mathematics test should be based on criteria that would ensure success in postsecondary study. Finally, the criteria, instruments, and overall evaluation program should be organized by the CMEC secretariat. The report proposed that a regular report to the Canadian people providing the collective achievement of Canadian students with respect to outlined criteria would also provide data to the jurisdictions and encourage increased accountability. This type of system evaluation program would provide an instrument for continuous improvement of local system performance and an improvement in the level of performance of Canadian schools would assure the well being of the country and of Canadians. They warned that without instruments to assess the educational systems, there was no way to tell how Canada's education system was performing (CMEC, 1988c).

The summary of the findings of the survey of the jurisdictions and the subsequent critical analysis were published exclusively for the CMEC in reports entitled "An Overview of Provincially Sponsored Student Achievement and System Evaluation Programs in Canada," (1988c) and "A Critical Analysis of Provincially Sponsored Student Achievement and System Evaluation Programs Examined from a Canada-wide Perspective" (1988a), respectively. These reports were compiled by a member of the staff of the CMEC secretariat under the direction of the ACDME Subcommittee on Priorities and Programs - ESE.

Becoming a Priority

According to the minutes of the 53rd CMEC meeting held in Regina, on September 26, 1988, national education indicators became a priority in the work plan of the Council. The minutes report that the indicators discussed were participation rates, standards and assessment.

At the fall meeting (September 28 and 29, 1987) the Council had agreed to establish a Steering Committee on Canadian Education Statistics. It consisted of deputy ministers who were appointed by CMEC to review, for improvement, the national education statistics system and make suggestions. The Steering Committee worked collaboratively with all the member jurisdictions to identify appropriate common education statistics. The Steering Committee formed two support committees: the Technical Committee on Elementary and Secondary Education and the Technical Committee for Postsecondary Education.[11] Each Technical Committee compiled specific indicator statistics for its area of education and prepared and reviewed the then draft protocol between the CMEC and Statistics Canada. The Steering Committee prepared a report for the ACDME on August 17, 1988.

[11] The Technical Committee for Postsecondary Education was disbanded in February 1988 and a new committee, the Ministerial Postsecondary Committee, was formed on June 7, 1988, to deal with postsecondary education statistics.

The work of Statistics Canada on education statistics was acknowledged and the August 17, 1988 Report of the Steering Committee on Canadian Education Statistics was discussed. It was agreed that the work of Statistics Canada on education statistics should continue. However the concept of national education indicators was considered to be additional information to that from education statistics.

The education indicators program was outlined by the chair of the ACDME Subcommittee on Priorities and Programs - ESE, the Deputy Minister of Education, Alberta, his second year in this position. The Subcommittee's outline was a proposal which included education statistics in collaboration with Statistics Canada, determination of expectations for achievement and the administration of an assessment instrument. The Council expressed specific interest in pursuing the idea of determining the provincial levels of expectations of secondary students and the literacy and numeracy achievement of 13-year-old students across Canada. Budget allocations were discussed for this program. Although there was no evidence in these minutes, it became apparent that at this meeting, Alberta agreed to act as the lead province and took the responsibility of turning this priority into an action plan. Also discussed at this meeting was the current work of the CERI on international indicators.

Another similar item on the agenda at this meeting was the then newly proposed Canada Scholarship Program. In its proposal, the federal Ministry of State for Science and Technology suggested that the scholars be selected by their success on a national test. All provinces were in agreement that the current means of determining scholars by each province was effective. Accordingly then, they believed that a national test was not required to assist with this since this was perceived to be federal intrusion into provincial matters.

When asked to comment on the proceedings of this meeting, the former chair of the ACDME Subcommittee explained:

> . . . again "the feds had done something." We've got to stop them. . . Finally I raised the issue of doing something that was productive, something that would

> help Canada assess itself. We needed to do something that would enable us to determine whether or not our students were performing as they were expected to perform. I had made that same suggestion a year earlier but it died. It died because the Chair at that time didn't quite know what we were talking about. So anyway, this year we came back with a proposal. Alberta Education prepared an outline of a proposal for a SAIP-like project for the subcommittee that I chaired. Fundamentally what we said was, what we wanted to do was to develop a test which would test children against what we believed they should know. This time when the CMEC made it a priority, I said great! Alberta will do it!

The national education indicators program was announced in a press release. In a follow-up news article to this press release, the November 7, 1988 issue of the Alberta Report noted that although national testing was on the horizon, this program was not about the standardized testing of every 13 year-old in Canada. The article quotes Olive Elliot, the then Edmonton Journal columnist, "Why *not* have standardized testing?" She further suggested having these tests in math and science as well as the suggested literacy test and publishing and comparing results. From this it appears that some members of the education community were ready for and welcomed national achievement indicators.

The Action Plan

In the fall of 1988, Alberta Education prepared a draft proposal for an education indicators project, for discussion at the December 15, 1988 meeting of the ACDME Subcommittee on Priorities and Programs - ESE. In briefing notes for the Chair of this ACDME Subcommittee prior to the December 15 meeting, the activities of the Steering Committee on Canadian Education Statistics were reviewed.

The January 17, 1989, minutes of the CMEC Technical Committee on Canadian Education Statistics - ESE indicated the education statistics currently proposed for study. Discussed was the possibility of a labour experience survey of students who had not

completed high school.[12] It was compared to the graduate survey and the tourism survey. Also discussed was the pilot of the Statistics Canada literacy survey, which was based on the Princeton National Assessment of Educational Progress (NAEP) survey. The literacy survey was to determine the performance of a sample of the population of Canada, aged 16 to 69, on selected literacy and numeracy problems.[13]

The agenda of the 47th ACDME meeting held in Toronto, (January 19 and 20, 1989) placed the discussion of the plan for the Canadian Education Indicators Project as presented by Alberta's Director of Student Evaluation and Records, on behalf of the ACDME Subcommittee on Priorities and Programs - ESE, first, and the Deputy Ministers' Steering Committee on Canadian Education Statistics, second. Prior to this meeting, members of the ACDME received a copy of the indicators proposal which had been revised by the Subcommittee on December 15, 1988. Minutes from the 47th ACDME meeting indicated that these two items were dealt with together. It was reported that a draft protocol between CMEC and Statistics Canada had been prepared and was ready for review by provinces prior to presentation to the Ministers in February, 1989. In reporting on the Canadian Education Indicators Project, minutes from this meeting indicated that the plan for this project was in keeping with the Ministers' intentions announced at their 53rd meeting and that, for the Canadian Education Indicators Project to be successful, the Subcommittee recommended that the work of the Steering Committee be expanded to include the Canadian Education Statistics program. As presented to the ACDME by Alberta Education, the goal of the CMEC Education Indicators Project was "To provide a Canadian information base that will enable provincial ministries to assess the performance of their education programs in comparison with Canada-wide standards" and the objectives were:

---

[12]This was conducted in 1991 by Statistics Canada. The findings were published in a report entitled School Leavers Survey.

[13]This was conducted in 1989 by Statistics Canada. The findings were published in a report entitled Survey of Literacy Skills Used in Daily Activities.

i. To enable the Council of Ministers to release an annual report showing the results achieved within each province relative to specific education indicators.

ii. To assist each ministry in being more accountable for the expenditure of public funds.

iii. T· ·ist each ministry in ensuring a high quality education for young Canadians preparing them to become responsible citizens and contributing members of society (Alberta Education, 1988 December, p. 1).

The proposal included a description, time line, and action plan for the three envisioned phases. Phase I ' 'ed the establishment of an indicators project, a frame of reference for the project, ar. mparison of expectations for achievement based on curricular documents. Phase II consisted of working with the Canadian Education Statistics project to monitor graduates' behavior after secondary school. Phase III consisted of preparing, administering and reporting on the results of a literacy and numeracy assessment of 13 to 15 year old students. The time frame was to begin in February 1989 and to finish in November 1990. This proposed start time explains why the SAIP was officially recorded to have started in 1989. Leadership in establishing a structure for the CMEC indicators project was written into the proposal as coming from Alberta Education (Alberta Education, 1988).

According to the ACDME minutes, the draft protocol for the Canadian Education Statistics Project; the goal, objectives, leadership, and action plan for the Canadian Education Indicators Project; and the expan.ion of the Steering Committee were agreed to, in principle, by the Deputy Ministers and consequently were placed on the agenda for the Ministers' 54th Council meeting, a month later.

CMEC's approval

At the 54th CMEC meeting he:d in Toronto on February 20 and 21, 1989, the Ministers approved the goal, objectives and start up costs of the action plan for education

indicators which was developed by Alberta Education under the mandate of the ACDME Subcommittee on Priorities and Programs - ESE. The presentations to the CMEC were made by personnel from the Student Evaluation Branch, Alberta Education. As part of their presentation, the representatives from Alberta used the 1988 CMEC internal report entitled "An Overview of Provincially Sponsored Student Achievement and System Evaluation Programs in Canada," to illustrate that while provincial testing occurs, it cannot be compared from jurisdiction to jurisdiction. In the CMEC document, entitled "A Critical Analysis of the Provincially Sponsored Student Achievement and System Evaluation Programs in Canada" similar conclusions are reached. A recommendation in this critical analysis was that existing evaluation programs be calibrated to more broadly based Canadian standards. The Alberta proposal suggested going further; the development of an indicators system specific to all of Canada. After much discussion about accountability, public funds, student achievement and priorities in education, the objectives i and ii were revised by the Ministers as follows:

> i. To enable the Council of Ministers of Education, Canada to release an annual report to the Canadian public showing the results achieved within each province relative to specific education indicators; and
>
> ii. To assist each ministry/department to evaluate student progress and to identify priorities in education (Alberta Education, 1989 February).

Reservations were expressed about the proposed public opinion poll in the second phase of the project. It was agreed that the project plan would be revised to compile existing public poll results instead.

The minutes of this 54th Council meeting indicated that the Canadian Education Statistics Project, the Canada Scholarship Program and the Council's Objectives also received equal discussion. The then draft "Protocol for Agreements Between the Government of Canada and the Provincial Governments for the Establishment of a Canadian Education Statistics Program" (1989) found agreement. The CMEC also

approved the proposal that the Deputy Ministers' Steering Committee on Canadian Education Statistics revise its membership and expand its mandate to oversee the management, but not the merger, of the Canadian Education Statistics and Canadian Education Indicators projects. The name of the latter project was considered to be easily confused with the Canadian Education Statistics project so there was considerable discussion about an alternative name for the Canadian Education Indicators. No decision was made at this meeting about a new name. Again, the federal Minister of State for Science and Technology made a purposeful presentation advocating a nationally administered examination on scientific literacy to assist in determining the candidates for the Canada Scholarships. The development and administration of this examination was emphatically rejected by the Council. The Council believed that the jurisdictions had suitable selection procedures and they questioned the appropriateness of such a request.

The Technical Committee on the CMEC SAIP

Demonstrating its new mandate, the second half of the March 22, 1989 meeting of the Deputy Ministers' Steering Committee on Canadian Education Statistics, in Toronto, dealt with the CMEC Education Indicators Project. Presentations were by Alberta's Deputy Minister of Education, Assistant Deputy Minister of Education and Director of Student Evaluation and Records. By agreement, the education indicators project name was changed to the CMEC School Achievement Indicators Project (SAIP). The revisions recommended by the 54th CMEC meeting were discussed and incorporated. A memorandum written by a representative from the Ministère de l Éducation, Québec on indicators and evaluating learning outcomes was also discussed. A discussion about distinguishing education statistics from the evaluation of learning outcomes and school achievement also ensued. The members agreed that education statistics should be distinguished from evaluation of learning outcomes and school achievement or indicators. The memorandum and the discussion served to inform all members of the Steering Committee about education

indicators. Also at this meeting, an organizational structure was developed for the SAIP. Operating under the direction of the Steering Committee, two Technical Committees or working groups, one already in existence, the Technical Committee on Canadian Education Statistics Elementary and Secondary and the other, the Technical Committee on the CMEC School Achievement Indicators Project, were formally established. (See Figure 4-2.)

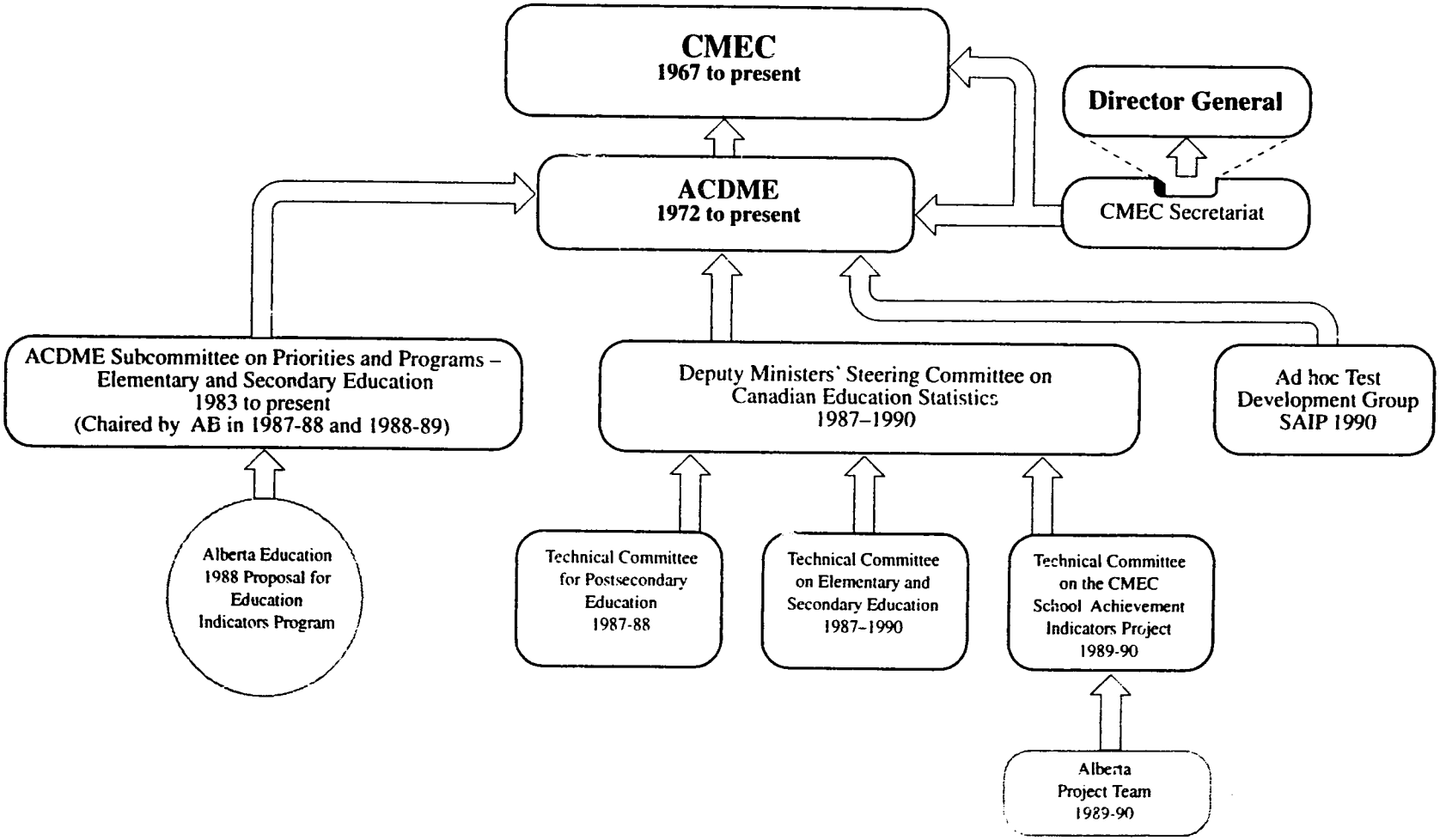
The first part of the minutes reported that the Protocol between Statistics Canada and the jurisdictions was ready for signing by the chair of the CMEC and the Minister responsible for Statistics Canada. This Protocol was implemented by the Canadian Education Statistics Council which convened its first meeting at the same time as the next ACDME meeting.

The Establishment of the SAIP

The first meeting of the Technical Committee on the CMEC School Achievement Indicators was held in Toronto on June 1 and 2, 1989. The purpose of this meeting was to have all members of the Technical Committee develop a common understanding of the SAIP, discuss implementation issues, and make recommendations to provide future direction for the Alberta Project Team. Documents such as the descriptive framework and methodology papers were provided as background materials. Fifteen members formed the membership of the Technical Committee on SAIP. There was one representative from each of the 10 provinces and two territories and three members from the CMEC secretariat. The representatives from the jurisdictions were Directors of Evaluation or experts in student assessment and evaluation.

In preparation for this meeting the Alberta Project Team developed or compiled the following documents: SAIP Descriptive Framework, Phase I Methodology Paper, Phase III Methodology Paper, Developing Flow Indicators by the Ministère de l'Éducation,

Figure 4-2 CMEC Organization for the SAIP 1987–1990



Québec[14] ; "An Overview of Provincially Sponsored Student Achievement and System Evaluation Programs in Canada" (1988c), and "A Critical Analysis of An Overview of Provincially Sponsored Student Achievement and System Evaluation Programs in Cana " (1988a).

In May 1989, all members of the Technical Committee on SAIP and the Alberta Project team had received a document outlining Québec's position on the proposal for an indicators project submitted by Alberta. In the document, Québec recognized the need for and the importance of establishing a Canadian information base to allow jurisdictions to assess the performance of their school programs, but recommended separating the statistical indicators, such as participation rates, retention rates and graduation rates, from the school achievement indicators and dividing the work between the two Technical Committees: the Technical Committee on Canadian Education Statistics in Canada - Elementary and Secondary and the Technical Committee on SAIP. Essentially Québec called for a separation of statistical indicators from school achievement indicators. Québec opposed the use of opinion polls and also believed that the comparison of expectations for achievement based on curricular documents and the development of assessment instruments were t ambitious and should be simplified. Although supporting it in principle, Québec believed that the entire project was too ambitious.

The chair of the Technical Committee (SAIP) was the Director of Student Evaluation and Records, Alberta Education. The Director General of the CMEC secretariat, presented the background of the project, explaining that student assessment and education systems evaluation had been discussed since 1987 at the level of the CMEC; explaining the Council's role in authorizing the survey of the provincially sponsored student achievement and system evaluation; and explaining the terms of reference of the Technical Committee on SAIP. The members of the Technical Committee were asked to cooperate fully with the

[14]This paper presented an explanation of how to calculate graduation rates by using the longitudinal, transverse or approximate methods.

Alberta Project Team in order to complete the first cycle of work on the SAIP over the next two years. Accordingly, consensus and free and open discussion were recommended by the chair.

The project leader from the Alberta Project Team presented the background and rationale for SAIP using the summaries and recommendations published in the CMEC reports entitled "An Overview of Provincially Sponsored Student Achievement and System Evaluation Programs in Canada" (1988c) and "A Critical Analysis of Provincially Sponsored Student Achievement and System Evaluation Programs Examined from a Canada-wide Perspective" (1988a). Also, the CERI international Indicators of Education Systems Project report submitted by the Directeur des statistiques et études quantitatives, Ministère de l'Éducation, Québec and the Director of Student Evaluation and Records, Alberta Education was discussed. These three documents were explained to be background for the CMEC SAIP. The first cycle of work on the SAIP was to determine the nature and scope of the project. The Alberta Project Team shared the Descriptive Framework for the SAIP. The Descriptive Framework outlined the purpose for the SAIP, the definitions of indicators, and the three proposed phases (Alberta Education, 1989a May).

Phase I focused on the involvement indicators: participation, retention and graduation rates and curricular expectations. The members of the Technical Committee on SAIP accepted the involvement indicators as defined in the Phase I Methodology paper. They also pointed out to the Alberta Project Team that the collection of this information was the mandate of the Technical Committee on ESE on Canadian Education Statistics; hence this Technical Committee would be asked to supply these statistics as well as enrollment, gender, and age information. The recommendations on curricular expectations raised many concerns, however the members of the Technical Committee on SAIP agreed to a feasibility study to determine the costs and the time involved.

Phase II included the identification of the public's expectations for and satisfaction with student achievement and graduates' expectation for and satisfaction with their school

experience. The Technical Committee on SAIP agreed in principle on the use of polls to determine this information and a look at existing polls was recommended. In the February proposal, the analysis of graduation and labor market success rates was part of this phase but it was dropped since the Project Team felt that it would be difficult to reliably link the information to other indicators. The Technical Committee on SAIP noted that a School Leavers Survey was planned to be conducted by Statistics Canada. Hence, this could provide information about graduates' expectations of and satisfaction with their school experience. As part of the public's expectations for and satisfaction with student achievement, a standard setting component was recommended by the Technical Committee on SAIP, prior to the administration of the test. Generally, then, the Technical Committee found agreement on phase two of the project.

Phase III included a literacy and numeracy assessment of grade nine and graduating students and was considered to be the student achievement indicator. Grade nine and the grade of graduation were preferred by the Project Team so as to link with other indicators. This would allow common curriculum to be tested and for comparison of expectations and satisfaction test results, curricular expectations and achievement results. Age, instead of grade, was brought up as a possible target population. The Project Team argued that if age was the criterion, then a common curriculum would not be ensured hence invalidating the expectations, satisfaction, and test result comparison. The grade versus age discussion was intense and inconclusive so an Ad Hoc Subcommittee on the Level of Testing of the Technical Committee on SAIP was set up. Representatives from Alberta, British Columbia, Manitoba, Ontario, Québec, the Northwest Territories and the CMEC secretariat were the members of this subcommittee struck to discuss and resolve this issue. When the assessment itself was discussed, the Technical Committee on SAIP suggested that the first focus should be literacy and mathematics and then perhaps science and culture. The Technical Committee on SAIP asked the Project Team to explore in a methodology paper the possibility of testing cultural literacy.

The Project Team explained the relationships among the indictors. Involvement indicators were related to curricular expectations which were related to student achievement. Involvement indicators were also related to graduate expectations and satisfaction and public expectations and satisfaction indicators. Public expectations and satisfaction indicators were also related to student achievement indicators. They also suggested that this system of indicators, now called a project, should eventually become a long-term program.

The Methodology Papers for Phases I and III received intense discussion. The methodology for Phase I was explained. It involved devising a rating scale to obtain numerical data when comparing the curricula of the jurisdictions for literacy and numeracy at the grade nine and graduating years. The Technical Committee on SAIP agreed to a feasibility study to determine this comparison. The source of the statistics (participation, retention and graduation rates) would be those generated by consulting with the Canadian Education Statistics Council according to their new Protocol. The transverse approach of determining involvement indicators as described by the Ministère de l'Éducation, Québec was recommended. The methodology for Phase III was explained to include the assessment of grade nine students (as this represents the year or next to the last year of compulsory schooling) and students who are in their graduating year on their literacy and numeracy skills using an interim instrument like the Canadian Adult Achievement Test (CAAT). The project team noted that a test specific to this project could be produced using test development procedures over three years. Consensus was reached on testing literacy and numeracy but not much else. A decision as to which instrument was going to be used was to be made on July 17, 1989. In addition, the Technical Committee on SAIP recommended that an annual report of the SAIP be published and further, that time lines and possible private funding by the Walter Gordon Foundation be considered.

Other items for discussion included the Steering Committee's approval of the Protocol between Statistics Canada and the Provincial Governments for the Establishment

of a Canadian Education Statistics Program; the March 21, 1989 C. R. Bronfman Foundation meeting (Alberta, CMEC and C. R. Bronfman) about their proposed national test in social studies.

The Protocol and the Canadian Education Statistics Program were established to determine comparable education statistics for the jurisdictions such as defining what full time equivalent students were and determining cost data for educating students. The chair of the Technical Committee on SAIP explained the management structure of the SAIP. Although the mandate of the Technical Committee on Canadian Education Statistics in Canada - ESE was quite different than the Technical Committee on SAIP, the two committees were managed by the Deputy Ministers' Steering Committee on Statistics. This management umbrella served to keep the two projects together but distinct thereby "keeping the statisticians at arms length" (Deputy Minister, Alberta Education, 1996) and not in control of the indicators project.

At the conclusion of this consensus-style meeting, an action plan was drawn up and a list of things to be completed before or by the next meeting was made. The second meeting of the Technical Committee on SAIP was scheduled for November 16 and 17, 1989. The age versus grade discussion was tabled until the June 22, 1989 meeting. The members of the Technical Committee were to review, critique, and provide a provincial position for or against the selection of the CAAT or any other commercial instrument. This review was to be sent to Alberta by July 10, 1989, so that an interim test instrument could be chosen on or by July 17, 1989. A published test was considered to be interim until such time that a test specifically designed for the SAIP could be constructed. Also, the members of the Technical Committee on SAIP were to provide descriptions of their respective provincial indicator projects to the Alberta Project Team for placement in Descriptive Framework of SAIP before June 22, 1989. Using the input from members of the Technical Committee on SAIP, the Alberta Project Team was to revise the Descriptive Framework and forward it to the ACDME meeting of July 31 to August 1, 1989.

Age versus Grade Testing

The Ad Hoc Subcommittee on the Level of Testing of the Technical Committee on SAIP met in Edmonton, on June 22, 1989. The minutes reported that representatives from all volunteer jurisdictions were in attendance. This Subcommittee was chaired by the Assistant Deputy Minister from Alberta Education. The agenda was to resolve the age versus grade proposal. The member from British Columbia requested that, if time permitted, the selection of the test instrument be discussed. The committee agreed. A spreadsheet of alternatives for level of testing by age, by curricular level, or by grade with underlying assumptions, type of information available, and data collection problems were provided as a discussion base by the Alberta Project Team. After almost an entire day of deliberations, the consensus decision reached was to test by age and at two critical levels. Testing by age would capture the range of educational programs offered across Canada and would be in keeping with the proposal originally approved by the CMEC, since the CMEC had made testing 13 and 15 year old students a priority in September 1988. Testing by age (13 and 16) was the resulting recommendation of this Ad Hoc subcommittee to the Technical Committee on SAIP. This decision was a significant change to the framework of the SAIP as proposed by the Alberta Project Team. According to the initial proposal, students in their graduating year should be tested to allow for meaningful interrelationship of indicators. In the time remaining a discussion ensued about the assessment instrument. This group also reiterated that since time was of the essence, using the CAAT would be reasonable as an interim test, but it was this subcommittee's position that it was necessary to develop a test. The chair concluded this matter by stating that the testing instrument would be discussed at meetings by teleconference of all members of the Technical Committee on SAIP on July 17, 1989.

The next meeting of the Technical Committee on SAIP was accomplished by two teleconferences, on July 17, 1989; one involving western and the other involving eastern jurisdictions. The Ad Hoc subcommittee's decision about testing by age was discussed.

Personal notes from this meeting reported that the Technical Committee on SAIP felt that testing had to be tied to curricular expectations as reflected in the modal grade for each age group. After considerable discussion and airing of provincial preferences, testing by ages 13 and 16 was agreed to by consensus by the members of the Technical Committee on SAIP. The next item on the agenda was the nature of the test. Again, after a long discussion where each representative shared the views of their jurisdiction, the Technical Committee on SAIP decided that the best test would be one that was specially developed for this program. To facilitate this, the Technical Committee on SAIP recommended that the timeline for administration be extended by one year to accommodate its construction. The Alberta Project team recommended the CAAT, a commercial test; however the Technical Committee on SAIP found questions in the CAAT unsuitable and ill-suited to the purpose of SAIP. The Technical Committee suggested that its use might cause embarrassment and public condemnation. The Alberta Project Team's recommendation about using the CAAT was problematic since it was an adult test and likely too complex for the 13 year old students. The Technical Committee also discussed compiling the results of the International Association for the Evaluation of Educational Achievement and the International Assessment of Educational Progress assessments and using these results as the achievement indicator. This was deemed to be unacceptable since not all provinces participated and they were not completely consistent with the objectives of the SAIP. The Technical Committee on SAIP's decision, then, was to recommend that a test be developed and that the timelines for the project be adjusted by a one year extension to pilot in 1990 and then test in 1991.

The Technical Committee noted that satisfaction polls had been conducted in the past but had not been related to public and graduate expectations. As a result, the Technical Committee on SAIP decided that two interrelated surveys would be recommended. Funding from the Gordon Foundation was discussed and involvement of the Foundation approved in principle by the Technical Committee as was the C. R. Bronfman

Foundation's offer of cooperation with their cultural literacy test. The Project Team explained the communication plan they proposed and the Technical Committee on SAIP strongly endorsed it. The decisions and recommendations of the Technical Committee on SAIP would be shared with the ACDME at their July 31 and August 1, 1989 meeting for their approval.

According to the minutes, the primary topic of the 48th ACDME meeting held on July 31 through August 1, 1989, in Toronto, was the CMEC SAIP. The minutes reported that the February 20, 1989 CMEC approval was discussed. The Chair of the Technical Committee on SAIP reported to the ACDME by describing the project as it had evolved thus far discussing the opinion polls, the target population of 13- and 16-year-olds, the test instrument, and the role of a charitable foundation. He explained the concept of the indicators system as being able to test the effectiveness of an education system; by determining if outcomes such as student achievement match student and public expectations and satisfactions.

The Phase I involvement indicators and curricular expectations indicators were presented as information. Specifically, that the involvement indicators were to be collected by the Technical Committee on Canadian Education Statistics; and although the curricular expectations indicators had received much support by the Technical Committee on SAIP and would yield useful information, the Technical Committee on SAIP had planned a feasibility study. It was mentioned that the curricular expectations indicators were also being studied by the OECD's CERI project.

The Phase II public and graduate expectations and satisfaction surveys were agreed to in principle by the Technical Committee, therefore, they recommended to the ACDME to approve the implementation of two surveys to determine the congruence between expectations of graduates and members of the general public and how well the schools satisfy parents, employers and postsecondary educators. However, it was noted that the Québec Minister of Education had been opposed to this type of survey in the past. It was

reported in the minutes that likely, Québec would agree if the survey expectations were related to outcomes and the survey was conducted at arms length to CMEC. Ontario expressed interest in school leavers and employment histories and in linking these to economic indicators. Additionally, Newfoundland called for a teacher survey on education. The ACDME approved the development of a public expectations and a graduate expectations questionnaire; and suggested that they be submitted to the Deputy Ministers' Steering Committee on Canadian Statistics for preliminary approval.

For the Phase III Achievement Testing, the Technical Committee recommended to the ACDME that the SAIP undertake testing by age, specifically, 13 and 16-year-old students. Thirteen represents an age where almost all students are still in the system and 16 represents the students in the system beyond the compulsory age. The two reference points allow for measurement in change or growth in achievement of students between the ages 13 and 16. The ACDME approved the testing of students at ages 13 and 16. The Technical Committee on SAIP recommended to the ACDME that should a commercial test not be feasible, then a test based on common curricular expectations among provinces and territories be developed. If this were the case then the pilot test would be scheduled for the spring of 1990 and the administration for the spring of 1991 to accommodate the development. This would represent a delay from the originally scheduled administration date of November 1990. However, the ACDME recommended to the Technical Committee on SAIP to choose a commercial test, administer it according to the original timelines, and evaluate its implementation.

The Technical Committee on SAIP reported to the ACDME that the Gordon Foundation was interested in funding the CMEC SAIP. The Technical Committee on SAIP recommended to the ACDME that a follow up on this funding offer be done. The ACDME recommended to the CMEC the involvement of the Gordon Foundation.

The C. R. Bronfman Foundation had shared with the Technical Committee on SAIP a cultural literacy test, for which it had produced a proposal, the Technical Committee

recommended to the ACDME to approve in principle cooperation with the C. R. Bronfman Foundation should its test be compatible with the purposes and objectives of the SAIP. The ACDME approved of this cooperation, in principle.

The Technical Committee on SAIP shared with the ACDME a series of communication strategies which included sharing information about SAIP with the jurisdictions, stakeholders, and the general public. It further suggested that members of the Technical Committee on SAIP serve as guests at conferences and meetings and that a communication package to prepared and sent to the jurisdictions. A short glossy brochure was planned to be prepared as information to the general public. ACDME approved this communication strategy and recommended that the Technical Committee on SAIP prepare documents to be submitted to the Deputy Ministers' Steering Committee on Canadian Statistics for preliminary approval.

Communication with Statistics Canada and the chair of the Technical Committee on SAIP occurred, by mail, on August 1, 1989. Statistics Canada shared its development of the school leavers survey and invited collaboration with the SAIP.

Data Collection for the Indicators

To help the Technical Committee on SAIP choose a commercial test, the Deputy Minister, Alberta Education "wrote to the Deputy Ministers of each of the other provinces and territories asking them to recommend an appropriate, readily available commercial test" (Alberta Education Briefing Notes, September 20, 1989). The survey and recommendation request was dated August 11, 1989. The former Deputy Minister reported "The only thing I was concerned about was that I did not want a standardized test, bought and applied. You see, I wanted to test what kids should know and what they needed to know as opposed to what they learned. So I sent out a questionnaire, but added one word to the end of it: OTHER, Of course everybody filled in OTHER because they didn't like any of the ones on the list."

On August 18, 1989, the Alberta Project Team and the chair of the Technical Committee on Canadian Education Statistics - ESE who represented British Columbia communicated about the availability of each piece of data required for involvement indicators. The chair of the Technical Committee on Canadian Education Statistics indicated that there would be difficulty in obtaining all necessary information for the involvement indicators for November 1989. The Alberta Project Team requested that the Technical Committee on Canadian Education Statistics survey the jurisdictions so as to determine when this information would be available.

On August 29, 1989, the chair of the Technical Committee on SAIP received a letter from the Director, Education, Culture and Tourism Division, of Statistics Canada. It outlined the work of Statistics Canada and the CMEC on the national program of education statistics. It referred to the Protocol, Statistics Canada publications for CMEC, and the review of classification systems used in educational statistics. Some of the work outlined included the transition of postsecondary graduates into the labour force, a survey of graduates and drop-outs of apprenticeship programs, regular annual surveys of enrollment and staff, a survey of private companies to determine the amount and type of training they provided for their employees and how much they invest in this, and a planned national survey on literacy skills for adult Canadians aged 16 to 69. He invited the chair of the Technical Committee on SAIP to contact Statistics Canada for assistance with the SAIP.

In September 1989, a delegation of Canadians, including the Alberta Deputy Minister of Education, the Alberta Director, Student Evaluation and Records Branch, (who was also the chair of the Technical Committee on SAIP) the Executive Director of Program Effectiveness from the British Columbia Ministry of Education, the Directeur des Etudes économiques et démographiques from the Québec Ministère de l'Éducation, the Statistics Canada Director of Education, Culture and Tourism, and the Coordinator ESE from the CMEC secretariat attended the General Assembly of the international Indicators of Education Systems Project, in Semmering, Austria (CMEC, 1991). The Alberta Deputy

Minister of Education, presented a speech to the representatives of the OECD countries. He spoke of the year or so in the development of a national indicators system in Canada. He described the three phases or components of the SAIP as being integrated and interrelated. He said that the student involvement indicators would be related to the student achievement indicator and this indicator related to the curricular expectations indicator and the public expectations indicator, hence the SAIP was described as an indicators system (Alberta Education, 1989). The General Assembly of the international Indicators of Education Systems Project, in Semmering, Austria, marked the end of the first phase of the international indicators project; the feasibility study.

Educan and the C. R. Bronfman Foundation wrote to the Director of Student Evaluation and Records Branch, Alberta Education, on September 22, 1989 requesting items that would be suitable for a cultural literacy test. Likely this request was sent to every Director of Student Evaluation in Canada. They notified Alberta that a request for confirmation of participation in the national cultural literacy test was sent to each Deputy Minister on September 18, 19[illegible]

In preparation for the 55th CM[illegible]ting in Toronto, briefing notes to the Minister of Education, Alberta, addressed [illegible]urrent issues of the SAIP. The ACDME approval and agreement of testing by age and the ACDME recommendation concerning selection of a commercially produced test. The briefing notes indicated the Alberta Project Team's position of using CAAT and Ontario's and Nouveau-Brunswick's agreement and the disagreement of using the CAAT by all other provinces and territories, according to the results of the August 11, 1989 survey. The recommendation of the Technical Committee on SAIP was to develop a test collaboratively and not use a commercial test even in the interim. A summary called Test Preferences by Provinces and Territories was prepared for the 55th CMEC meeting. No consensus was evident among the jurisdiction responses. The majority of the jurisdictions were in favor of developing a test specific to the SAIP, either from scratch or from provincial achievement examination questions. Alberta's official

position was to support the idea of developing a test specific to the SAIP. The briefing notes recommended Alberta's Minister of Education to urge the ACDME to reconsider its position on the commercial test and to promote test development.

According to the minutes of the 55th CMEC meeting held in Toronto on September 25, 1989, the SAIP was the second item on the agenda. A progress report was presented by the Deputy Ministers' Steering Committee on Canadian Education Statistics. The latest revisions as requested by the ACDME were approved, except that a special work group, the Ad Hoc Test Development Group, was given the mandate to determine the feasibility of developing a test specific to the SAIP. In this subtle way, the minutes of the CMEC meeting showed that the Ministers approved of developing an instrument, not using a commercial one. The Ad Hoc Test Development Group was to report to the Deputy Ministers' Steering Committee on Canadian Education Statistics, who then reported to the ACDME. It was noted in the minutes that the development of the SAIP was proceeding well and was due to the direction and major contribution of Alberta. Concern, however, was expressed about involving all stakeholders in education and reservations about the public opinion polls were also made. A delay was suggested on these polls. The CMEC gave approval to the CMEC secretariat to contact the Gordon Foundation for funding. The Deputy Minister of Education, Alberta reported on the presentation he made to the international Indicators of Education Systems General Assembly held in Semmering, Austria on September 19 and 20, 1989.

On September 26, 1989, an official news release on behalf of the CMEC and Statistics Canada announced the signing of the Protocol and the establishment of a Canadian Education Statistics Council. The Protocol outlines collaborative work between the federal government and the provinces on the national education statistics program. This announcement closed the 55th CMEC meeting. In a follow-up communiqué, CMEC also described the SAIP. The graduate and public expectation and satisfaction polls and the curricular expectation indicators were missing in its description.

On September 26, 1989, a discussion paper on literacy, written by a member of the Alberta Project Team, was circulated to all members of the Technical Committee on SAIP. A conceptual and theoretical framework of literacy was outlined and three key recommendations were presented. These were that the term 'literacy' would connote the purposeful use and understanding of written language in a variety of situations; that reading and writing be involved in the test in both the multiple choice and written format; and that the test be developed with the input of educators from all provinces and territories. (Hochachka, 1989) Written on the copy of this paper in the personal notes of the Alberta Project Team member was this comment: "If there's tension now, it'll be even thicker when this is out," illustrating that the collaborative work and consensus decision making process suggested by the terms of reference of the Technical Committee on SAIP was extremely difficult.

Further communication (September 28, 1989) with the chair of the Technical Committee on the Canadian Education Statistics program indicated that the information on the involvement statistics would not be available until the 1990-1991 school year.

As a result of the breakfast meeting of the Deputy Ministers' Steering Committee meeting of September 26, 1989, correspondence (October 3, 1989) was set in motion to establish a special committee, the Ad Hoc Test Development Group, to prepare a report on the practicality, feasibility, and implications of developing the test. This report was due for the January 16, 1990 ACDME meeting. Representatives from British Columbia, Alberta, Ontario, Québec, Prince Edward Island and 3 members from the CMEC secretariat volunteered to make up the Group. The chair was the Assistant Deputy Minister of Education, Alberta. The Group met on November 14 and 15, 1989.

On October 5, 1989 the President of Educan received a letter from the Director-General of the CMEC indicating that ongoing cooperation with the Technical Committee on SAIP could be expected on the cultural literacy test and that provincial responses to their September 18, 1989 request for jurisdiction participation were likely forthcoming.

On October 16, 1989, the President of Educan received a response from Alberta to the proposed cultural literacy test, the "Heritage Test." Many suggestions were of[illegible]ed. However, the most important suggestions were the inclusion of questions testing higher order thinking skills and that the testing ages should be 13 and 16 years old so that the assessment would be in keeping with the proposed SAIP assessments.

On October 17, 1989 a teleconference with all members of the Technical Committee was held so as to share the decisions of the ACDME and the CMEC meetings in August and September. The decision to develop the test was announced. A special committee, the Consortium for Developing the SAIP Assessment Instruments was set up to construct the assessment instruments. The C. R. Bronfman Heritage Test was discussed. Written as notes on the teleconference agenda was this comment, "Ensure that they don't see this as a full endorsement. We'd have to see the actual test before we could approve liaison" indicating how important some members believed the SAIP assessment instrument to be. The framework for the literacy assessment with its key recommendations, the framework for the numeracy assessment, and the statistics on involvement indicators were also discussed.

In preparation for the November 16 and 17, 1989, meeting of the Technical Committee, the Alberta Project Team prepared several detailed reports, strategies, and methodology papers. A report on the involvement indicators identifying when, and what data the jurisdictions could supply was written. A feasibility study to devise a strategy for studying and comparing curricular expectations across the jurisdictions was undertaken. A methodology paper was written that described the use of existing polls to collect data on public expectations of and satisfaction with the education system. A strategy for developing test instruments to assess literacy and numeracy at ages 13 and 16 was devised. A prototype of the CMEC indicators report was developed. The implementations timeline for the project was revised, a communications plan was developed, and the information package was written. All of this work was completed in a five month time span by a six

member team from the Department of Education, Alberta, who had other duties and responsibilities.

The Technical Committee on Canadian Education Statistics met on November 9 and 10, 1989, partly to discuss the data collecting procedures and information requirements for the involvement indicators of the SAIP. Correspondence indicated that a discussion ensued on a series of possible definitions of the terms: participation, retention and graduation. The Technical Committee on Canadian Education Statistics discussed concerns such as that data collection on graduation was a new requirement for some jurisdictions.

The Ad Hoc Test Development Group met on November 14 and 15, 1989, in Toronto. The major decision resulting from this meeting was to hire a university professor from a university in Ontario to draft a Request for Proposals. This Request for Proposals was to determine the contractor for construction of the assessment instruments.

The Technical Committee on SAIP met on November 16 and 17, 1989, in Toronto. Major items discussed were Canadian participation in the Semmering, Austria, CERI General Assembly, the report from the Ad Hoc Test Development Group, the framework for a numeracy assessment, the communications strategy, the definitions of involvement indicators, and methodology papers on the opinion poll, [illegible]dy, and the curricular expectations comparison. Major papers on the four latter topics were prepared by the Alberta Project Team Each paper clearly outlined each indicator and provided recommendations for collecting the appropriate information. The proposed School Leavers Survey of Statistics Canada was analyzed for Phase II of the SAIP. However, there was considerable lack of fit between its probable design and the design required to determine expectations and satisfaction of graduates.

The Ad Hoc Test Development Group reported that it contracted out the drafting of a Request for Proposals for the construction of a test for the SAIP. The Request for Proposals would accompany the Group's report to the ACDME planned for January 1990. The timeline from the initial awarding of the contract to the first full testing was revised

from May 1990 to May 1992. The members of the Group shared their views on how to involve teachers, strategies to develop tests, and how to determine the level of achievement in literacy and numeracy. They suggested that it would be best to establish what literacy and numeracy are by representing their competencies on a continuum that described five stages. The assessment instrument would be developed to determine the level of student achievement along the five level continuum. Members of the Technical Committee expressed reservations about the suggested hierarchy of learning within the proposed levels of literacy and numeracy.

The members of the Technical Committee from Newfoundland led a discussion based on a Mathematical Literacy paper. Although deliberations ensued, it was agreed that each member of the Technical Committee would submit written comments to the Chair of this committee by January 4, 1990, and that the paper would be revised accordingly. It would then be made available to the successful candidate of the Request for Proposals and form the basis of the mathematics assessment.

An Alberta Project Team member presented the proposed communications strategy. After much discussion, the Technical Committee noted that both a long and a short term communications strategy was necessary. The members also pointed out that jurisdictional and national responsibilities and costs should be publicized. In light of the discussion, it was agreed that revisions should be made prior to submission to the Deputy Ministers' Steering Committee and the ACDME.

A member of the CMEC secretariat reported on the state of the international Indicators of Education Systems Project. Alberta and Québec were commended to their contributions to this project to date. Phase Two of this project was planned to begin 1990 and jurisdictions were encouraged to express interest.

The chair of the Technical Committee on Canadian Education Statistics - ESE reported on the data-collection strategies for the involvement indicators. Definitions of the various indicators were discussed, as was the helpful role of personnel with Statistics

Canada. It was agreed to contact the then newly formed Canadian Education Statistics Council for advice and recommendations. The first report on these indicators, using September 1990 data, was published in June 1992 entitled A Statistical Portrait of Elementary and Secondary Education in Canada.[15]

Another Alberta Project Team member presented the curricular expectations paper. The feasibility study was reported to be underway and its results would be available before the March 1 and 2, 1990 meeting. The Technical Committee recommended that the study take into account that curricular expectations would likely be a part of the construction of the assessment instrument. Also discussed was the public satisfaction survey-feasibility study. The Technical Committee determined that the intent of this survey was to find out the public satisfaction with respect to the curricular expectations. It was noted that not all jurisdictions supported this indicator. A feasibility study would be drawn up for consideration of the Technical Committee at the March 1 and 2 meeting. As well, the graduate expectations survey was defined as a series of questions asked of students just prior to graduation, those one year after graduation, and those five years after graduation. The Technical Committee requested a detailed proposal that included costs and required resources for the August 1990 ACDME meeting.

The Technical Committee considered a draft response written by the chair, to the C. R. Bronfman Foundation, providing feedback to this foundation about its proposed cultural literacy test. Reservations were expressed and the Director-General of the CMEC indicated that the jurisdictions could relate to the C. R. Bronfman Foundation bilaterally.[16]

After this Technical Committee meeting, the Alberta Project Team resumed their work on the revisions arising from recommendations suggested by various members of the Technical Committee meeting. The review-feedback on the mathematics discussion paper

---

[15]The Canadian Education Statistics Council published an earlier edition in 1990, however this edition only included data on one indicator.

[16]No jurisdiction entered into a bilateral agreement with the C.R. Bronfman Foundation and plans for its proposed cultural literacy test were subsequently discontinued.

was received and the paper was updated for the Deputy Ministers' Steering Committee meeting of December 6, 1989. The curricular expectations rating scale was updated and a field study completed for final decision to proceed or to drop at the March 29 and 30, 1990 meeting. A prototype public expectations study was developed for the Deputy Ministers' Steering Committee meeting and a complete plan prepared for the Technical Committee for a final decision to proceed or not. Terminology that hinted of it being a poll was removed. The substance of this study was curricular expectations hence allowing relationships to be drawn to other indicators. The communication strategy was revised in line with the Technical Committee recommendations for approval at the January 16, 1990 ACDME meeting. Finally, the graduates' expectations study was updated for approval by the Technical Committee's March 29 and 30 meeting.

According to the minutes of the 49th ACDME meeting held January 16, 1990, in Toronto, statistics and indicators dominated the agenda. There were reports from the Deputy Ministers' Steering Committee, Ad Hoc Test Development Group, the Technical Committees on the SAIP[17] and the Elementary and Secondary on Canadian Education Statistics, the Canadian Education Statistics Council, and about the OECD Indicators Project. The Ad Hoc Tests Development Group reviewed the Request for Proposals. The ACDME recommended that the CMEC be involved and take a leadership role in constructing the specific test, but that they be developed by consultants. The ACDME recommended that the short term communication strategy be accepted with the exception of the idea of the press conference proposed for February 1990. The ACDME also recommended that the CMEC hire a fuli-time coordinator for the SAIP and establish a SAIP advisory body. The Canadian Education Statistics Council reported that its first publication entitled A Statistical Portrait of Elementary and Secondary Education in Canada was almost ready for publication.[18]

[17]In the minutes of this meeting the name of the School Achievement Indicators Project was changed to the School Achievement Indicators Program; no explanation or discussion was provided.
[18]This was published in August 1990. It included information on the graduation indicator, only.

At the 56th CMEC meeting held on February 13, 1990 the Council examined the recommendations of the ACDME on the SAIP and agreed to all of them. The mandate was given to the CMEC secretariat to name a coordinator for the SAIP and to the ACDME to establish a SAIP advisory committee. The SAIP advisory committee would consist of deputy ministers and would serve to provide direction to all working groups of the SAIP.

The Technical Committee on SAIP met on March 29 and 30, 1990, in Toronto. According to the minutes, the major items discussed were the activities of the Alberta Project Team on Phase II of the SAIP, the overall schedule of the SAIP, and the role of standards in the SAIP. Also provided as information were the recent CMEC decisions concerning the SAIP, the Request for Proposals, and the Canadian participation in the CERI international Indicators of Education Systems project. The Canadian participation in the CERI international Indicators of Education Systems project involved collaboration with other OECD countries to identify indicators associated with finance and student flows. This group planned to meet in Paris in May of this same year.

The Alberta Project team presented an overall schedule for the SAIP. This included that a preliminary report on Phase I could be published in 1990 and a full report could likely be published in 1992. The representative from Québec noted that the three indicators proposed for Phase II had not yet been approved by the ACDME. He asked members of the Technical Committee if these large scale studies were even a realistic undertaking. He stressed that it would be preferable to concentrate on Phase III the achievement and Phase I the involvement indicators, as his Minister of Education was not interested in the indicators proposed by Phase II. A lengthy discussion ensued. Finally, the Director-General of the CMEC pointed out that it was the responsibility of the Deputy Ministers' Steering Committee, the ACDME, and the CMEC to make those decisions. He reminded the Technical Committee of its responsibility to set up clear options for the Steering Committee on the feasibility, resource requirements and relevance of Phase II indicators.

Each component of the Phase II indicators was discussed first separately and then as a whole. The curricular expectations feasibility study recommended the study of a set of curricular expectations to serve as a frame of reference for comparing curricula across the jurisdictions. A rating approach was not favoured especially if it ended up ranking the jurisdictional curricula. Members stressed that a descriptive analysis of curricula would be more favoured as the basis for comparison among jurisdictions. Also, members of the Technical Committee generally believed that the main purpose of this component was to support the construction of the assessment instruments. The public's expectations and satisfaction study was presented by the members of the Technical Committee from British Columbia and Saskatchewan. They indicated that this study would help determine appropriate standards by consulting with the public and this indicator would allow for the interrelationships of the indicators. Again, the Technical Committee was reminded that this indicator was not supported by some Ministers of Education. Although no consensus was reached, the Technical Committee suggested that perhaps some information could be gathered about public expectations and satisfaction, even after the testing had occurred. Next, the graduate expectations and satisfaction study was discussed. Similarly, no consensus for a recommendation to the Deputy Ministers' Steering Committee could be reached. Objections raised were the cost, the appropriateness, the types of diplomas and or certificates that represent a graduate, and that the survey was too general. Then, Phase II of the SAIP as a whole was discussed. No consensus was reached on the need for its inclusion in an indicators program. That these indicators would provide the process and context of the input and output indicators was largely ignored by the members of the Technical Committee. A subcommittee was scheduled to meet in Montreal on May 17, 1990 to deal with this issue and then report to the Technical Committee meeting planned for June 14 and 15, 1990, in Québec City.

The Technical Committee broke into three groups to discuss the role of standards in the SAIP. One group discussed that four types of standards may be involved in the SAIP:

results over time, comparisons with other provinces, norms for Canada and absolute standards. The second group discussed the advantages and disadvantages of standards. Without standards, it was noted that the mean will become the standard. And the third group discussed the need for effective standards. The Technical Committee agreed that the topic of standards for the SAIP was crucial and required further ongoing study.

The Technical Committee on SAIP met on June 14 and 15, 1990, in Québec City. This was to be the last meeting of the Technical Committee on SAIP. According to the minutes, the major items discussed were Phase II, the expectations indicators, including a report from the May 17, 1990 subcommittee report on this component of the SAIP, and the discussion of levels of achievement for the Phase III of the SAIP. Also discussed was the report from the Ad Hoc Test Development Group, the communication strategy, and the issues related to reporting results.

On the discussion of the Phase II component of the SAIP, the minutes reported the recommendations of the May 17, 1990 subcommittee. Although the members of the subcommittee recognized the need for the various components of a system of indicators and understood that reporting and interpretation of indicators would be clearer with the expectations and satisfactions indicators, they, collectively, did not favor the development of a curricular expectations indicator, nor a student (graduate) expectation and satisfaction indicator, nor a parent (public) expectation and satisfaction indicator. Instead, they recommended that the Phase II component of the SAIP contain a curriculum expectations description of similar documents from the provinces and territories. This would provide contextual information for the assessments. Also, the subcommittee recommended that an ethnographic survey be made of students, parents, and teachers participating in the assessment. This would provide contextual information for the subsequent report. The subcommittee did not see a need to obtain further information from other members of society, therefore no variation of the public expectations and satisfactions indicators were proposed. Although consensus was not reached on all aspects of this information

gathering, the Technical Committee agreed to gathering descriptive information rather than numerical ratings on these expectations and satisfaction indicators.

On issues arising in the reporting of indicator results much of the discussion centered on data to be presented in the reports on involvement and achievement. A mock-up involvement report was presented. A long discussion ensued on the definitions of participation, retention, and graduation rates. Decisions were made as to numerators and denominators to be used in the calculations, for example, the numerator for the participation of students age 15 would be, the number enrolled in secondary school, over the denominator of all 15-year-olds available for enrollment. These decisions were provided to the Canadian Education Statistics Council. The members of the Technical Committee also discussed the concept of reporting on the level of student achievement. Although, both the literacy and numeracy discussion papers explained the concept of determining the level of student achievement by determining the degree of competency that students have, Committee members questioned this practice. Reference was made to the use of levels by Statistics Canada in the Survey of Literacy Skills Used in Daily Activities. The Technical Committee recommended that the achievement report include an explanation on this assessment strategy and eventually consensus was reached on this issue.

On June 20, 1990 a meeting was held to announce the SAIP to Nongovernmental Organizations (NGOs). Representatives from the Canadian Association of School Administrators, the Canadian Chamber of Commerce, the Canadian Education Association, the Canadian Teachers' Federation, and the Conference Board of Canada were invited to attend an all day information meeting about the SAIP. The agenda included the formal announcement speech of the SAIP by the Alberta Minister of Education, a description and current state of the SAIP by the Assistant Deputy Minister of Education, also from Alberta, and a question and answer series from the participants. According to the Minister's speech the reasons for introducing a national indicators program included

> increasing interest in the question of standards in education. We need concrete evidence of what we are achieving in schools because that makes good educational sense. You can't aim for quality if you have no way of assessing progress. You can't decide what to change and what to maintain in the schools if you don't know what works and what doesn't. . . . Hence this nation-wide school achievement indicators project.

Similarly, the 1989-90 CMEC Annual Report explains: "The CMEC School Achievement Indicators Program will help Canadians know how well we are doing in educating our youth" (CMEC, 1991, p.9).

The Management Structure for the SAIP

At the 50th ACDME meeting held August 14, 1990 in Toronto, once again the SAIP occupied the majority of the meeting. The ACDME received reports from the June 14 and 15, 1990 meeting of the Technical Committee and the June 20, 1990, NGOs meeting. They provided direction on the management strategy for the numeracy and literacy assessment, the expectations indicators, the collection of the data for the involvement indicators, and the creation of the SAIP advisory committee. It was reported that three responses to the Request for Proposals[19] were received but that none was acceptable. The ACDME recommended the appointment of a program director, two assistants, and teams of experts to develop, validate, and administer the test instruments. A selection committee was established to set up this management organization. On the involvement indicators, the ACDME delimited them to the curricular expectations. For these, the ACDME recommended that focus groups be contacted and that jurisdictional curricular expectations be incorporated into the definitions, the criteria, and the actual tests. Also recommended was that background be gathered on the students and teachers in a questionnaire form at the

[19]The three proposals varied so greatly in their estimated cost and description of the anticipated work that it was evident that none of them was capable of undertaking the endeavour.

time of the assessment. The deputy ministers agreed to provide data for the involvement indicators to the Canadian Education Statistics Council, with a view of reporting in 1991. The ACDME recommended that the Ad Hoc Test Development Group, the Deputy Ministers' Steering Committee on Canadian Education Statistics and the SAIP, and the Technical Committee on SAIP all be disbanded. According to the mandate given to them at the 56th CMEC meeting, they created one SAIP advisory committee. It was chaired by a deputy minister[20] and composed of jurisdictional representatives who had been members of the Technical Committee on SAIP. It was known as the SAIP Policy Advisory Committee.

At the 57th CMEC meeting held in Saint John on September 24, 1990, the Council made a decision on the mechanism for the development of the assessment instruments. Since no suitable contractor was available, the ministers revised the ACDME's management strategy to accommodate a consortium approach to test development. Alberta was identified as the lead province; a position to be shared with Québec. As a result, the Consortium for Developing the SAIP Assessment Instruments was set up. The Consortium consisted of representatives from Alberta, Québec, and the CMEC secretariat. The chair of the Consortium was the Assistant Deputy Minister of Education, Alberta. The Consortium was made up of two different Assessment Instrument Development Teams, one that had the responsibility for the mathematics assessment and the other that had the responsibility for the reading and writing assessment. Each Development Team had two representatives from each province in the Consortium. The organizational structure for the SAIP also included the newly formed body of Provincial Contacts for SAIP Assessment Instrument Development. Representatives from each jurisdiction were named as contacts and they were to provide advice to the Assessment Instrument Development Teams.

On December 3, 1990, a special meeting of the CMEC was held in Ottawa. The Minister of Education from Ontario raised many concerns about the SAIP. Ontario was concerned about the representativeness of the Development Teams, the criteria, the

[20] The first chair was the Deputy Minister from British Columbia.

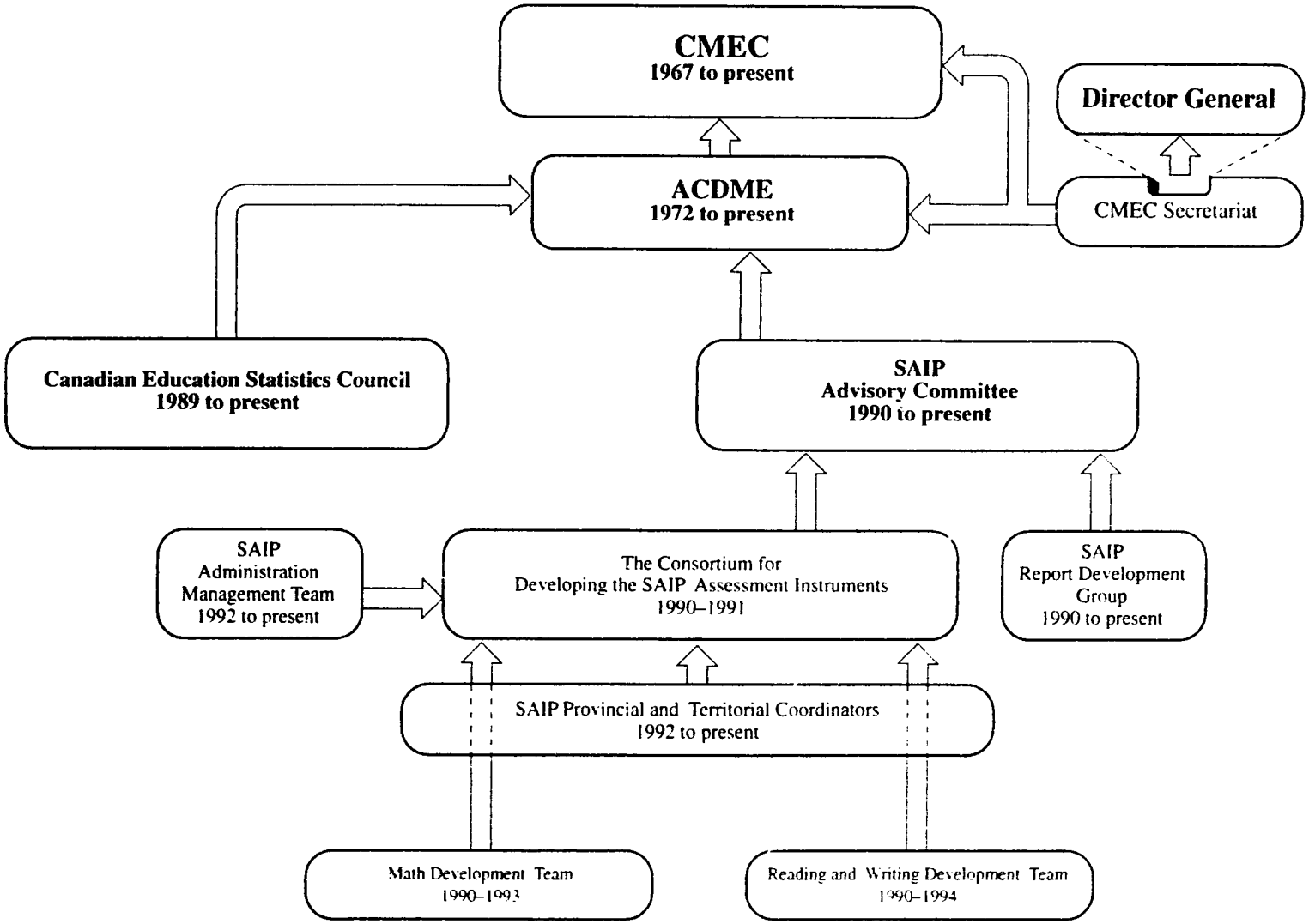
sampling procedures, and the reporting procedures for the assessments. To deal with these concerns, Ontario suggested that a Memorandum of Understanding be developed and agreed to by all jurisdictions, at the next CMEC meeting.

At the 51st ACDME meeting held January 21, 1991 in Victoria, reports were heard from the SAIP Coordinator and the chair of the Consortium. The Assessment Instrument Development Teams presented preliminary criteria for the literacy and numeracy assessments to the Provincial Contacts for SAIP Assessment Instrument Development on January 16, 1991. Decisions were made using consensus. The chair of the Consortium stated he was satisfied with the progress of the Development Teams and that management by consortium was desirable for a pan-Canadian program such as the SAIP. The Memorandum of Understanding requested by Ontario was discussed. The ACDME gave Ontario the mandate to draft the Memorandum of Understanding for approval of the SAIP Policy Advisory Committee. The ACDME also agreed to instruct the SAIP Policy Advisory Committee to prepare a report on SAIP for the next CMEC meeting.

At the 58th CMEC Toronto on February 25, 1991, the Memorandum of Unde v Ontario and dated February 1991 was discussed in camer ced by the Council so Ontario opted out of participating in th status. This was a concern to the CMEC since slightly mor king students in Canada are educated in Ontario.

At the 52nd ACDME meeting held August 8, 1991 in Toronto, a progress report on the instrument development was heard and the management structure for the SAIP was completed with the addition of the SAIP Report Development Group. (See Figure 4-3.) The SAIP Report Development Group consisted of deputy ministers, directors of evaluation or named representatives, one from each jurisdiction. The ACDME set up the Terms of Reference of the Report Development Group. The Report Development Group was to make recommendations to the newly formed SAIP Policy Advisory Committee, on

Figure 4-3 CMEC Organization for the SAIP 1990–1993



pertinent options for reporting on the SAIP. Reporting to the public on the achievement of students at one of five levels was the first reporting issue with which the Group had to deal.

In describing the interprovincial work that had gone on during the 1990-91 year, the CMEC Annual Report indicates:

> Over the year, CMEC acquired considerable expertise in learning how to manage a complex project that required sharing resources, contracting out specific tasks while retaining overall control, clarifying numerous issues, adapting the program in response to feedback, and putting in place a sound financing strategy. The entire process has enhanced the Council's resourcefulness and affirmed its role in bringing the provinces and territories together to address national issues in education (CMEC, 1992, p.6).

As mentioned previously in this study, as the SAIP developed so did the CMEC in its ability to manage and lead a large-scale project.

At the 59th CMEC meeting held in Calgary on September 23 and 24, 1991 the Memorandum of Understanding of the SAIP was discussed again, this time in camera. A new Memorandum of Understanding due for December 9 and 10, 1991 was mandated. The new Memorandum of Understanding was to include Ontario in the Consortium for Developing the SAIP Assessment Instruments and therefore as part of the Assessment Instrument Development Teams, to define literacy and numeracy and base their criteria on curricula, outline the report's use and dissemination, and develop the budget and the timeline for field testing and administering the assessment instruments.

At the 53rd ACDME meeting held December 9, 1991 in Toronto, the 1991 SAIP Memorandum of Understanding was readied for CMEC consideration.

At the 60th CMEC meeting held in Toronto on December 9 and 10, 1991 the 1991 SAIP Memorandum of Understanding was approved by 11 of the 12 jurisdictions of Canada. Saskatchewan abstained from approval of the Memorandum of Understanding and

agreed only to observer status in the SAIP. Ontario was now a part of the Consortium and the Development Teams. The management structure for the SAIP was now in place.

In the 1991-92 CMEC An Report the Director General explained:

> This has been a long, difficult, and at times frustrating, process, quite without precedent in Canadian education but by year's end we were well on track, having successfully field-tested the assessment instruments[21] in June. There was an encouraging feeling that the road ahead was now clear of serious obstacles and that the SAIP would proceed as planned (CMEC, 1993, p.5).

Two years and two months had elapsed since the SAIP was recognized a priority of the CMEC, received approval, and then finally came into existence. The time, in terms of a lifetime, is considered to be short, however to take two years just to frame a project is indeed a "long" time.

The Science Assessment

Over the next two years the assessment instruments were developed for administration in spring 1993. Because of some difficulties in selecting parallel reading passages in both French and English, the reading and writing assessment instruments were not ready for administration until the spring of 1994. To exemplify the process of instrument development for this national indicators program, the historical development of the SAIP science assessment will be explained.

Conception and Approval

According to records kept by Alberta Education, a proposal was put forward by members of Alberta Education to include a science assessment with the SAIP at the 58th ACDME meeting held August 10, 1993. This proposal included a description of two two-

[21]The Director General was referring to the field testing of the mathematics assessment in the spring of 1992.

hour components to the assessment: a two hour test which would have multiple-choice, short-answer, and extended-response questions; and a two hour practical that would have the students do 10 tasks in a laboratory-like setup. A skeleton budget was included. Alberta was named as the lead province in the proposal. According to the minutes of the ACDME, the deputy ministers considered the idea of a science assessment. The ACDME agreed that the inclusion of a science assessment was in keeping with the newly proposed expansion of the SAIP and CMEC and Statistics Canada program into a pan-Canadian Indicators program of educational success. Whether or not the Alberta Education proposal was actually presented to the ACDME is not clear because the Science Development Team leader commented that ". . .what had happened is they had some indication who the province was that would have taken the lead, but that didn't follow through, and for awhile it sat in quandary as to whether they were going to get the leadership on this. So at that point we put together a proposal and took it directly to the Ministers. That proposal was accepted and it went from there."

Preparations continued in Alberta to present this proposal to the 64th CMEC meeting held September 27 and 28, 1993, in Victoria, British Columbia. The proposal grew into four parts. Part one described the test design. There would be a two hour written component with multiple-choice, short-answer, and extended-response questions and a two hour practical component with a reduced number of 5 to 8 tasks (Shavelson, 1990). The students would be assessed on a five level scale according to the level of learning they demonstrated. The criteria for the assessment, although not yet developed, would be based on four areas: science knowledge; science and technology; nature of science; and science, technology and society. The development team would include a leader, four science specialists, a laboratory assistant and two word processors. This part of the proposal also briefly described the expected field testing, scoring, and final assessment.

Part two provided a literature review and was entitled "Scientific Literacy Achievement Assessment." According to the proposal's interpretation of the literature, a

description of scientific literacy was presented. Scientifically literate students were expected to be able to understand basic knowledge, understand scientific processes, evaluate evidence, apply science concepts, understand relationships between science and technology and the limitations of each, and understand the societal context for science. Essentially the four areas of the assessment were described.

Part three described the practical component of the assessment. The practical assessment would be based on analyzing the students' achievement in six skill areas: initiating and planning, collecting and recording, organizing and communicating, analyzing, connecting, synthesizing and integrating, and evaluating. Each of these areas was described.

Part four of the proposal was the action plan, timeline, and budget. A line-by-line breakdown of the requirements for development of a scientific literacy assessment, along with the time that it would take for completion of each task and the associated costs were presented. This proposed budget was considerably higher than the budget in the initial proposal. The start date was anticipated for August, 1993 and the completion was set for September 1996, with the administration of the assessment being April 1996. This proposal was developed by Alberta Education and distributed to all members of the CMEC on September 17, 1993 (Alberta Education, 1993).

At the 64th CMEC meeting held in Victoria on September 27 and 28, 1993, the need for national standards in education and national cooperation among the jurisdictions dominated the discussion. The Council agreed to the Joint Declaration: Future Directions for the Council of Ministers of Education, Canada. A national agenda in education was proposed in the Joint Declaration. The action plan of the national agenda in education included national consultations on education conferences[22], examination, comparison and

[22]The First National Consultation on Education took place in Montreal in May 1994. The Second National Consultation on Education was held two years later in Edmonton in May 1996.

joint curriculum development[23], continuation of the SAIP with the addition of a science assessment,[24] the implementation of the Pan-Canadian Education Indicators Program (PCEIP)[25], the encouragement of open and distance learning projects, and closer communication with business and other relevant education partners. Hence, in two short meetings the SAIP science assessment was conceived and approved. What it would be based on and how it would be developed is yet to be explained.

Between the 64th CMEC meeting and the meeting in December 13 and 14, 1993 to plan the development of the science assessment, the secretariat of the CMEC took over the coordination of the initiative. A November survey of all the jurisdictions was made to determine the names of each jurisdiction's representative (their Science Contact) and to determine which jurisdictions were interested in forming the Consortium for the development of the assessment instrument. By December 6, 1993, Alberta and Ontario, mandated by their respective ministries, had expressed interest in being part of this consortium.

The Framework

December 13 and 14, 1993 was the first meeting of the Science Contacts. The four-part proposal was presented to the Science Contacts by Alberta. After an intense discussion, reported in the meeting minutes, the term scientific literacy was dropped from the descriptive title of the assessment. The idea of assessing student achievement in science at one of five levels was an issue as was the idea of assessing students at ages 13 and 16 as opposed to assessing students at specific grades. Again, after much discussion a consensus

[23]A Pan-Canadian Protocol for Collaboration on School Curriculum K-12 was adopted by the CMEC at the February 1995 meeting. The first collaboration began in July 1995 with the Pan-Canadian Framework of Goals and Outcomes for Science K-12.

[24]The continuation of the SAIP refers to the second cycle of assessment administration, for example the mathematics assessment is planned for the spring of 1997, the reading and writing for the spring of 1998 and the second administration of the science assessment is planned for the spring of 1999.

[25]A Statistical Portrait of Elementary and Secondary Education in Canada report was published in April 1996. It included data on student participation and student achievement indicators. These represent two of the six indicators of the PCEIP. The others, equitable accessibility to education, successful transitions between school and work, and public satisfaction with the education system are currently work in progress.

was reached to assess using levels and that the assessment be carried out on populations of 13 and 16-year-old students. Although the weighting of the practical was discussed, the idea of including a practical component to the science assessment was unanimously approved. The Science Contacts suggested regrouping the four areas to be assessed into three: knowledge of science, nature of science, and science, technology and society. Suggestions to include the assessment of attitudes and to collaborate with the then current TIMSS were also made. The contacts discussed finalizing the consortium membership in January, 1994, updating the Memorandum of Understanding (See Appendix F) to include the science assessment, and adding the requirement of ACDME and CMEC approval of the model of assessment by February 28, 1994.

According to Alberta records, Alberta incorporated the suggested changes and ideas and by January 10, 1994 a document entitled "CMEC School Achievement Indicators Program Science Assessment" that described the model of the assessment was written. On January 12, 1994, the secretariat of the CMEC sent this document to the Science Contacts for feedback. The recommendations and comments included that the term scientific literacy be avoided throughout the document, that a curriculum fit when the criteria are developed be ensured, that achievement standards be set, and that the difficulty of level five be considered. Again, Alberta incorporated the suggested changes into a revised document entitled "CMEC School Achievement Indicators Program Science Assessment" dated February 4, 1994.

By February 17, 1994, four provinces had named representatives to the consortium: Alberta, Saskatchewan, Ontario and francophone Nouveau-Brunswick. March 3 and 4, 1994, marked the first meeting of the Science Contacts of the Consortium. According to the meeting minutes, their Deputy Ministers had officially volunteered their province's participation in the consortium at the 59th ACDME meeting held on January 17,

1994[26] in Toronto. This first meeting was held to discuss the 1991 Memorandum of Understanding and suggest changes for the new Memorandum of Understanding, set the Consortium Agreement, discuss the February 4, 1994, document entitled "CMEC School Achievement Indicators Program Science Assessment," plan the work project, and the set overall budget for the project. At that point, the members of the consortium approved that Alberta would lead the Consortium in the development of the science assessment and that Nouveau-Brunswick would represent the francophone sector and produce the final assessment instruments in French. Because Saskatchewan had not been involved in the SAIP mathematics and SAIP reading and writing assessments, it was interested in taking part in the development and in doing an evaluation of whether or not it was appropriate for Saskatchewan to become involved in the SAIP Science Assessment of its students. The existing management structure was discussed (See Figure 4-4) and it was agreed that the CMEC secretariat would hire a SAIP National Coordinator to serve as overall manager of SAIP. The committees in existence were the SAIP mathematics assessment Report Development Group, the SAIP reading and writing assessment Administration Management Team, and the SAIP provincial and territorial coordinators (called Contacts) all of which were overseen by the SAIP Policy Advisory Committee.

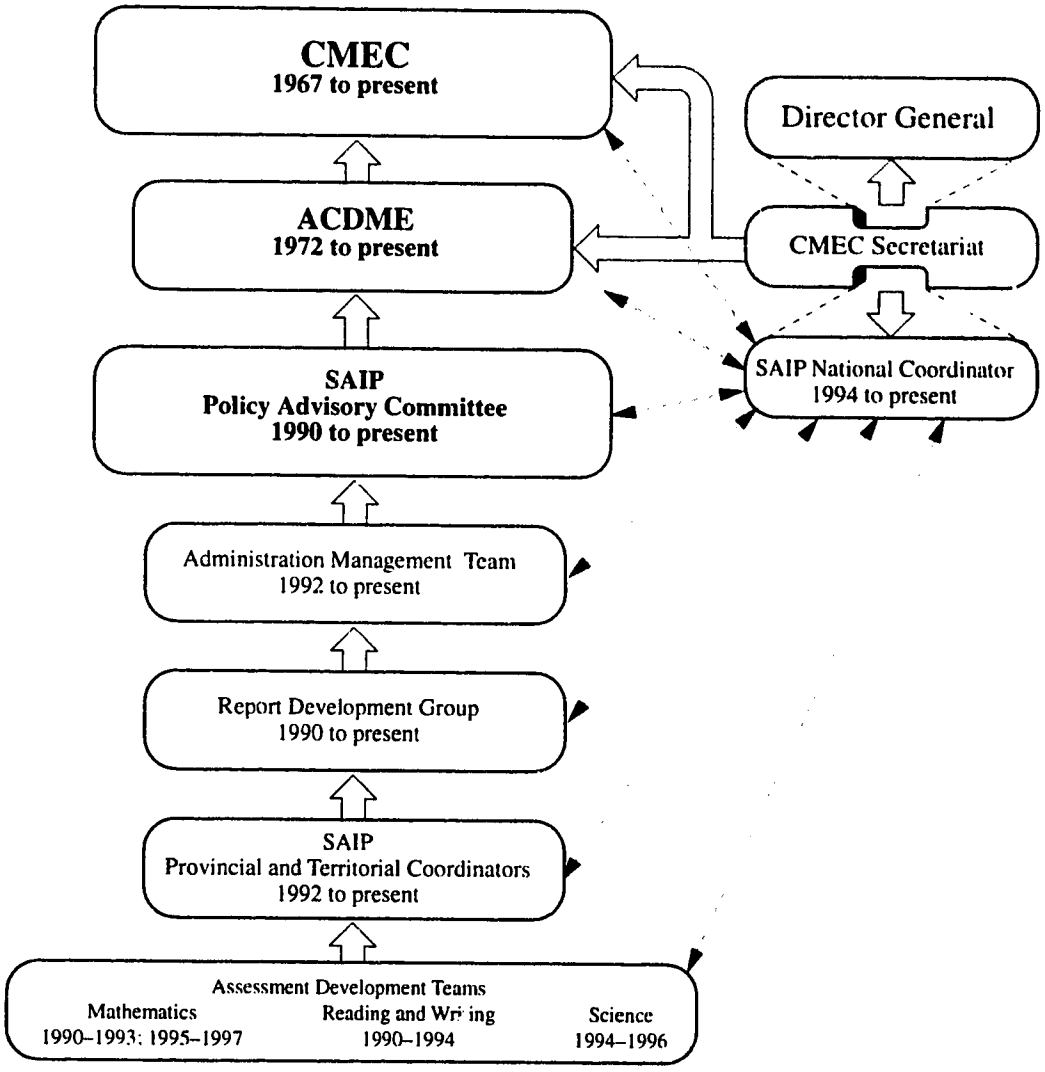
Also reported in the meeting minutes of March 3 and 4, 1994, was the discussion and inclusion in the Consortium Agreement of the duties of the members. The Consortium Agreement[27] described the scope and requirements of the assessment project, outlined the responsibilities of the individuals from their provinces who were to become members of the Science Development Team, described the caliber of individuals who would be chosen as members of the Science Development Team, listed the special responsibilities of each of the

---

[26]The minutes of the ACDME report that the ACDME approved in principle the draft model for the design of the science assessment.

[27]Appended to the Consortium Agreement was the 1994 Memorandum of Understanding, the May 20, 1994 Science Assessment Framework and Criteria, a Work Plan and Reporting Schedule, Biographies of the Science Development Team members, Budget, Computer Equipment and Software Inventory, and a Copyright and Confidentially Agreement.

Figure 4-4 CMEC Organization for the SAIP 1993–1996



provinces, outlined the financing and transfer of payment plan, outlined the computer and other equipment purchase and eventual disposition and maintained the ownership of all materials produced or purchased to be the property of the CMEC. The Consortium expanded the membership of the proposed Science Development Team from 4 Science Specialists to 6, from 2 Word Processors to 3 and added a Performance Skills Specialist. The Science Development Team actually consisted of a Development Team Leader, a Performance Skills Specialist, five Science Specialists, a Laboratory Clerk and three full

time equivalent Word Processors. The researcher was one of the teachers named to the Science Development Team. It should be noted that in keeping with the agreement's description of the caliber of the members of the Science Development Team, four of the five teachers named as Science Specialists had won major provincial or national awards. The Consortium Agreement commenced February 1, 1994 and terminated in September 1996.

Also at the March 3 and 4, 1994 meeting, the Consortium agreed to establish a Science Development Team that could function in both official languages, a mechanism to provide a pan-Canadian consultation on the criteria for the science assessment, a pan-Canadian curriculum analysis, a pan-Canadian item writing strategy, an instrument development and field testing strategy, a NGO and Expert Review of the frameworks and criteria for the assessment, and a mechanism for the validation of the levels and the items on the assessment.

According to minutes, the first meeting of the Science Development Team was held in Toronto, Ontario from March 21 through 25, 1994. Six of the eight members were present. The Consortium Agreement and the Memorandum of Understanding were discussed, and funding, the timelines, and the scope of the project were reviewed. The team discussed the February 4, 1994 document entitled "CMEC School Achievement Indicators Program Science Assessment." Revisions were discussed. The knowledge of science was then subdivided into four domains representing the disciplines of chemistry, biology, physics and earth science. An unpublished skills document from Alberta illustrating four levels of assessment in six skill areas was analyzed (Alberta Education, 1994 February). The team recommended that there be five levels of assessment in the six skills with the view to maintaining the similarity of the written and practical assessments. A plan was made to order the required computer equipment[28] which would enable participants located in different parts of the country to work on the assessment

[28]Team members has access to e-mail and faxes to facilitate communication.

instruments. The work plan for the next meeting and the division of labour were made. The changes to the framework would be made, the computers would be ordered, and reading would be done in preparation for the writing of the criteria for each of the domains of the assessment. The next meeting was scheduled in Fredericton, from April 18 through April 22, 1994.

Alberta revised the framework and the "CMEC, SAIP Science Assessment Framework" dated March 30, 1994, came into existence in both French and English. This document was mailed to the Science Contacts at all ministries of education for their approval and solicited comments by May 4, 1994. All of the replies were received and compiled by Alberta Education by May 16, 1994. An April 6, 1994, draft document entitled "Guide for the Construction of Questions" was written by team members from Alberta and Ontario. Although it had focused on setting up and organizing its operational features, the main function of the Science Development Team was to produce a draft framework to circulate to ministry representatives (Science Contacts) for consultations.

The Criteria

The minutes report that at the April 18 through 22, 1994 meeting of the Science Development Team, a member of the CMEC secretariat shared the April 14, 1994 draft version of the Consortium Agreement. Suggestions were made by the Science Development Team to the CMEC secretariat on wording, timelines, scoring, and the action plan for standard-setting. The team agreed that February 1, 1994 would be the start date and saw their role as coordinating the development of national standards which the provinces and territories would establish. These recommendations were taken by the CMEC secretariat and the Consortium Agreement was finalized.

However, the main work of this meeting was the writing of the criteria for each of the seven domains represented in the assessment. The concept of levels was already in place in the original framework as it matched both the mathematics, and the reading and

writing assessments. The criteria had to be completed and ready for the June 28 and 29, 1994 meeting in Ottawa. This was the review of the framework and criteria by NGOs, stakeholder organizations and experts agreed to at the March 3 and 4th, 1994 meeting of the Science Contacts. This was an absolute deadline. The design of the assessment, the framework, and the criteria had to be developed in a very, very short time. The decision was made to accept but enhance the framework that Alberta had developed. In developing the criteria per level for each domain, the team decided to use an already existing construct rather than to create a new one. The team looked at the Science Council of Canada, 1984, Science For All Canadians, Report 36, Science For All Americans, and an associated text Benchmarks for Science Literacy, and the framework for the TIMSS assessment and at the curriculum guides from all the jurisdictions. After much discussion and deliberation, the team decided to base the criteria on a modification of the work done by the American Association for the Advancement of Science, Project 2061, as presented in the Benchmarks for Science Literacy. As there were eight members in attendance (seven team members and the Science Contact from Saskatchewan), four teams of two were formed. The members of the team were matched according to their area of specialization. With this division of labor, the team wrote the criteria for the four knowledge domains of the assessment in two days. These were edited and revised by the entire team. Once completed, one day was spent on selecting, modifying, and developing the criteria for the nature of science and the science, technology and society domains of the assessment. However because of time constraints, this task was not completed in the allotted time and so each of the team members worked on the criteria for these domains on their own in preparation for the next meeting that was held in Edmonton, Alberta from May 2 through May 6, 1994.

Additional work was also done between these meetings. The translation of the criteria was completed by Nouveau-Brunswick, writing and refining the item writers' manual was done by Ontario and Alberta, and writing the performance item writers' guide was undertaken by Saskatchewan. All revisions, edits and additions/deletions of criteria

were then forwarded to Alberta for a first draft of the framework and criteria document. Alberta was ultimately responsible for developing the criteria for the five levels of each of the science skills.

At the third meeting of the Science Development Team held in Edmonton from May 2 through 6, 1994, the minutes reported that the criteria of the knowledge domains were polished and then, as a group, the team worked together on finalizing the criteria for the nature of science, science, technology and society, and the science inquiry skills. Again the team divided up the work and the criteria were read and revised to show uniformity. The product of this work was the May 6, 1994 "SAIP Science Assessment Criteria" document. Besides this major accomplishment, the Memorandum of Agreement was discussed and the Consortium Agreement was studied line by line and necessary changes were made. The June 26 through 28, 1994 and the June 29 through 30, 1994 meetings were planned. The item writing sessions dates and locations were chosen. The test design was discussed but not finalized. Suggestions were made to assess 13 and 16-year-olds to level four but assess 16-year-olds who have completed grade 10 to level five. Reporting was discussed but also was not finalized. Suggestions were made to report achievement on two of the four knowledge domains and on nature of science, science, technology and society for everyone or to report the overall achievement as an aggregate of all seven domains.

After the Edmonton meeting, the May 6, 1994 "SAIP Science Assessment Criteria" document was revised and edited by Alberta and redated May 10, 1994. This new document included reference to the use of the recommendations found in the Science for all Americans and the Benchmarks for Science Literacy. A decision was made to include the criteria with the framework and a new document entitled the "Council of Ministers of Education, Canada, School Achievement Indicators Program, Science Assessment Framework and Criteria" dated May 20, 1994, resulted. This document reflected the changes requested by the ministries as a result of their review of the March 30, 1994 version. (See reference to May 4, 1994 reply and May 16, 1994 compilation.) The May 20,

1994 document was sent to the ministries and to the science experts (invited to the June 29 and 30th, 1994 meetings) for consultation and comments and responses were requested by June 15, 1994. Most of the responses arrived in Alberta between June 9 and June 17, 1994. A line by line analysis had been completed by most ministries by this time.

The fourth team meeting was held in Ottawa from June 26 to 28, 1994, with six members in attendance. The team reviewed the ministerial responses to the May 20, 1994 "Council of Ministers of Education, Canada, School Achievement Indicators Program, Science Assessment Framework and Criteria" document. The concern was raised by some jurisdictions and team members that in some jurisdictions topics such as earth science and physics were not taught to the students of the targeted age groups. The consensus reached by the team was that the opportunity to learn should not dictate parameters of the criteria of a science assessment. Changes were made after two long days of discussion. The June 28, 1994 meeting ended at 3:00 a.m. June 29, 1994. The result was the "Council of Ministers of Education, Canada, School Achievement Indicators Program, Science Assessment Framework and Criteria," June 28 document. Item writing plans for the summer were finalized based on the replies sent from the ministries offering to host item writing sessions. The format and design of the assessment instruments were discussed but not finalized. The reporting was discussed and a consensus resulted that this be based on four categories: the knowledge; nature of science; science, technology and society; and on the inquiry skills, but was not finalized.

A national information and consultation half-day meeting was held in Ottawa on June 29, 1994 with NGOs and science experts and a full-day consultation meeting with science experts on June 30, 1994. In preparation for these meetings, the NGOs received the framework document and the science experts received information about the SAIP, the Report on the SAIP Mathematics Assessment (CMEC, 1993) and the framework and criteria for the science assessment. In the question and answer session and the small group discussions that occurred over these two days, numerous issues arose. Among them the

following questions surfaced: What was the rationale for the assessment of achievement at one of five levels? Why use age sampling when grade sampling is much more convenient? What sampling procedure will be used? Why not report on the achievement by each of the four knowledge domains? How will the validation be done? and, What about those students who did not have the opportunity to learn the subject matter? These were powerful questions. Some of these were the same questions that the Technical Committee had addressed in 1990, others were the same questions that the ACDME had addressed throughout the development of the SAIP and still others were the same questions that the development teams addressed as they prepared their instruments. The members of Science Development Team and the leader of the Science Development Team addressed these questions orally, on-site. Comments about the framework and criteria were considered and a new version of the "Council of Ministers of Education, Canada, School Achievement Indicators Program, Science Assessment Framework and Criteria," dated July 15, 1994 document was prepared.

On July 7, 1994, a Science Development Team teleconference meeting was held which involved the seven Science Development Team members, the National Coordinator, and the Science Contacts from Ontario and Saskatchewan. The purpose of the teleconference was to obtain a team consensus position on each of the questions (as summarized by the CMEC secretariat) posed by the experts and stakeholders at the NGO meeting. According to audio tapes, the two main questions addressed were the levels and ages used in the assessment. The historical background of the use of levels and ages in the SAIP assessment was shared and the fundamental basis of their use in the SAIP assessment was stressed. In attending to the concerns of the experts, it was decided that the concept of five levels in the nature of science domain might be tentatively applied; for all other domains, it would be applied. This was agreed to by the team. The consensus reached by the team members is reflected in the letter sent to the attendees of the NGO, dated Sept. 6, 1994.

On July 15, 1994, another meeting of the SAIP Science Contacts was held in Toronto, Ontario. An update of the activities of the Science Development Team was given, issues arising from the June 29 and 30th, 1994 meetings were discussed question-by-question, and the "Council of Ministers of Education, Canada, School Achievement Indicators Program, Science Assessment Framework and Criteria" document dated July 15, 1994 was reviewed. The Science Contacts recommended that the Science Development Team and the CMEC secretariat address these issues and reply to the attendees of the June 29 and 30th, 1994 meeting in an open letter format. This letter was written by the leader of the Science Development Team, with advice from former Technical Committee members and sent to the attendees on September 6, 1994. The Science Contacts made further suggestions about the "Council of Ministers of Education, Canada, School Achievement Indicators Program, Science Assessment Framework and Criteria," July 15, 1994 document. This document was revised and dated August 1, 1994 and mailed to all the Science Contacts for their final input and approval by August 10th, 1994. It was approved and forwarded to ACDME for official approval at their August 22 to 23, 1994 meeting. The model of assessment presented in this document was endorsed by the ACDME. The ACDME acknowledged the need to change the management model but extended the management model used for the previous assessments to the science assessment. Also, all provinces except Prince Edward Island agreed to participate in the assessment. Prince Edward Island elected observer status. This document was forwarded by the National Coordinator for the SAIP and the Science Development Team leader to the CMEC at the 66th meeting (September 26 and 27,1994) in Charlottetown for final approval. However, the Ministers did not deal with this agenda item. Although not formally approved, the criteria (and framework) had been a result of a series of consultations, feedback, and revisions by ministry representatives and was essentially the product of these ministries.

Instrument Development - Item Writing

All jurisdictions except Prince Edward Island had teachers and or consultants participating in the item writing process. Item writing sessions began July 11, 1994 and continued until September 1, 1994 and were held in nine different jurisdictions. The item writing was based on the "Council of Ministers of Education, Canada, School Achievement Indicators Program, Science Assessment Framework and Criteria" document dated July 15, 1994, however any items that were written on criteria that were subsequently deleted, were also deleted. Most of the practical tasks were written by teachers from Saskatchewan, British Columbia and Yukon. Items for the written component of the assessment were written in sets based on a pictorial scenario. The members of the Science Development Team revised these scenario sets of questions and the Word Processors in Alberta and Nouveau-Brunswick entered them into the computer. A computer program was used that allowed the members of the Science Development Team to choose questions by scenario. This facilitated the building of the assessment instruments.

Unofficial Field Test - 1994

In Fredericton, from September 12 until September 23, 1994, the seven members of the team took raw items and their scenarios and reworked them keeping the interesting, representative, and criteria-referenced items. Items were switched from one scenario to another and scenarios were used to build two "A" tests of 24 items, and three "B" and "C" tests of 54 items each. Each "A" test included four items from each of the six domains and were all of level three. Each "B" test had three items from each of the knowledge domains at levels one and two and one item for each knowledge domain from level three. It also had five items from the nature of science and the science, technology and society domain for levels one and two but three items for level three. Each "C" test had three items from each of the knowledge domains at levels four and five and one item for each knowledge domain from level three. It also had five items from the nature of science and the science,

technology and society domain for levels four and five but three items for level three. For each level of the knowledge domains, nine items were field-tested except level three where 10 items were field tested. For each level of the nature of science and the science, technology and society domains, 15 items were field tested except level three where 22 items were field-tested. For ease of field testing in one-hour sittings, each "B" and "C" test was split into two tests of approximately 27 items. During this process, the idea of the extended-response question was reduced to field testing a few two and three point questions. Most of the open-ended questions were one point questions. About one third of the questions field test d were open-ended questions. In the day and a half remaining of the two week meeting, two team members sorted the 30 practical tasks produced by the teachers to chose 20 of them to field test.

The two "A" tests, six "B" tests, six "C" tests as well as the 20 practical tasks were field tested in the four provinces of the consortium in the month of October 1994 in an unofficial field test. The "A" tests were tried in grades eight, nine, and 11 (science, biology, chemistry, and physics) classrooms. The "B" tests were tried in Grade 8 and 9 classrooms and the "C" tests were tried with grade 11 and with advanced science students. Ten practical tasks were tested in each province. The purpose of the unofficial field tests was to check the performance of the items, validate the levels (level one with grade eight and level five with the advanced science students) and have the ministries validate the items. The 14 tests and four practical tasks were mailed to the ministries for their review and feedback. Each item was criterion-referenced to the "Council of Ministers of Education, Canada, School Achievement Indicators Program, Science Assessment Framework and Criteria," document dated October 10, 1994. This document was revised once more since the August 1, 1994 revision to take into account the comments from the team and comments which had arrived over the summer. The ministries were provided feedback forms per test and were asked to comment on the curriculum and age alignment of

each item and the intrinsic value of each item. Their responses were requested for November 10, 1994.

The majority of the responses were received by November 17, 1994. A tally of each of the ministry's responses per item was made. Generally, the concepts for the items in the "A" test were taught by grades seven, eight, or nine; all the concepts for the items in the "B" test were taught by grade nine, and all the concepts for the items in the "C" test were taught by grade 11 or 12.

Official Field Test - 1995

Members of the Science Development Team scored their respective papers and summarized the results. Based on these summaries and the tallied comments from the ministries, items were chosen to build the tests for the Official Field Test. These field tests were built in Edmonton, from November 21 until 25, 1994. The team of seven agreed to discard those items which were unanimously criticized by the ministries and revise others according to these criticisms. Two "A" tests with 12 items, two per domain, at level three were built. (See Table 4-1.)

Table 4-1 The Design for Field Test A

Domain	Number of Level Three Questions
Biology	2
Chemistry	2
Physics	2
Earth Sciences	2
Nature of Science	2
Science, Technology, and Society	2
Total number of questions	**12**

Two "B" tests with 66 items were built. They consisted of four items at levels one and two for the four knowledge domains, except level three which had two items for each, and five items at levels one and two for the nature of science and the science, technology and society domains, except level three which had three items each. (See Table 4-2.)

Table 4-2 The Design for Field Test B

Domain	Number of Questions at Level One	Number of Questions at Level Two	Number of Questions at Level Three
Biology	4	4	2
Chemistry	4	4	2
Physics	4	4	2
Earth Science	4	4	2
Nature of Science	5	5	3
Science, Technology, and Society	5	5	3
Totals	**26**	**26**	**14**

Two "C" tests with 66 items were built. They consisted of four items at levels four and five for the four knowledge domains, except level three which had two items for each, and five items at levels four and five for the nature of science and the science, technology and society domains, except level three which had three items each. (See Table 4-3.)
The items at level three on the "B" and "C" tests were exactly the same so that for each knowledge domain there were eight items field tested per level and that for the nature of science and the science, technology and society domains, there were 10 items field tested per level. Slightly more than one-third of the questions were open-ended. However, it was decided that an open-ended question would have the same weight as a multiple choice

question. The idea of an extended response type of question was completely dropped[29]. By dividing the team into two groups the "A", "B" and "C" tests reached a skeleton completion

Table 4-3 The Design for Field Test C

Domain	Number of Questions at Level Three	Number of Questions at Level Four	Number of Questions at Level Five
Biology	2	4	4
Chemistry	2	4	4
Physics	2	4	4
Earth Science	2	4	4
Nature of Science	3	5	5
Science, Technology, and Society	3	5	5
Totals	**14**	**26**	**26**

in three days. The remainder of the week was used to discard any obviously weak practical tasks, refine the inquiry skills criteria to ensure that a different or higher level skill was required at each of the levels and choose the 13 practical tasks for the field tests.

At the end of the meeting, Alberta had the responsibility of the "A" tests and the one "D" test; Nouveau-Brunswick, the "B" tests and the Student Data Booklet; Ontario, the "C" tests; and Saskatchewan the 13 practical tasks. The one 60 item D test was compiled by choosing two representative items from each of the six domains at each of the five levels. This served as further validation of the levels and of the items. Immediately after the meeting in Edmonton, Saskatchewan had to order all the materials to do the national field tests in April and May of 1995. This was very difficult as ordering materials through the government purchasing agency was not very efficient timewise because a tendering process

[29]The weighting of an extended-response question was problematic psychometrically. Since each question was referenced to one criterion, with a multiweighted, multi-criteria question, the referencing would not be clear.

had to be used and the materials ordered were not necessarily the exact ones expected. From the time the tender was sent out until it was due, it took anywhere from two to four weeks to supply after they were notified. In some cases it was from five to seven weeks from the time the orders were placed and the time the materials came in, in late January 1995. These materials were required for the February 21 to 22, 1995 in-service of the Science Contacts. Other provinces had other equally pressing responsibilities. Nouveau-Brunswick translated the changes. Ontario prepared the questionnaires and, along with one member from Alberta, prepared the student score sheet specific to each test. The Science Development Team leader made preparations for the December 12 to 15, 1994 team meeting and the February 21 to 22, 1995 in-service of the Science Contacts.

At the December 12 to 15, 1994 Science Development Team meeting held in Toronto, the nature and the purpose of the field tests were discussed, the involvement of the provinces was planned, the preparations for the printing and mailout were made, the tracking of the "A", "D" and the practical tasks was accomplished by preparing numbered stickers in sets of three, the administration of the practical tasks that included an in-service was planned, the required statistical data was discussed, and the final versions of the "A", "B", "C", and "D", tests and practical tasks were established in both languages. Teacher and student comments were also requested.

At the 61st meeting of the ACDME held in Toronto on February 3, 1995 the "CMEC SAIP Science Assessment Framework and Criteria," document dated February 1, 1995 and the official field test materials were submitted for approval. However the ACDME minutes report only an update of the activities on the work of the Science Development Team.

According to the minutes of the 67th meeting of the CMEC held in Toronto on February 27 and 28, 1995, all documents relating to the SAIP science assessment were approved in principle, in preparation for the upcoming official field test. The National Education Agenda, however, dominated the discussion of this CMEC meeting.

The Science Contacts received an invitation to attend the February 21 and 22, 1995, in-service for the administration of the field test held in Toronto. Response was requested by January 18, 1995. All jurisdictions responded with a total of 25 participants. There was at least one representative from each ministry to as many as eight from one particular ministry. On February 21, 1995, the assessment instruments were shared, the field test strategies were explained, an overview of the field test was presented and the participants received the "Information for Field Test Administrators" handbook. An open question-and-answer session about the field test administration followed. On February 22, 1995 the participants were involved in a hands-on in-service of the practical tasks. The Science Contacts were provided with an average cost per practical task-kit as information to help them decide if their jurisdiction would be involved in the oversample. Special instructions about each task were shared and then participants worked through each task. The feeling in the air was very positive and the instruments were well received.

A team meeting was held back-to-back to this in-service session with all eight members of the team in attendance. This was the first meeting that included the Laboratory Assistant. The details of mail-outs of the written component and practical component were discussed, return of the students booklets to the Toronto marking centre was set for May 22, 1995, the scoring-planning meetings for the written and practical components were planned for June 5 to 12, 1995 and July 17 to 28, 1995, respectively, and the editing, printing and distribution of the bulletin and brochure were discussed. As well, a look ahead at the preparation of the official assessment instruments, and questionnaire (to gather data about the student's opportunity to learn, attitudes toward science and demographic information) and the sampling for the official administration was made.

In the spring of 1995, between April 15 and May 15, 1995, every jurisdiction was involved in the field testing. Each jurisdiction received information about sampling representative grade eight and 11 classes and each participated as much as it was able to from as few as 57 students, to as many as 783 students who wrote field tests in any one

jurisdiction. More than 3 500 students across Canada were involved in the field test and validation of the items and levels of the written assessment. The field testing of the "A", "B", and "C", tests validated the items. The repeat questions from level three on the "B" and "C" tests were used to validate level three and the "D" test, which was administered to both grade eight and 11 students, was used to validate the levels for each of the domains, but level three in particular. Also, by administering the instruments to students of all ability levels at grade eight and 11, levels one, two, four and five were validated. Each jurisdiction field tested seven different practical tasks of the 13 for their validation. This involved approximately 1 500 students across Canada. Approximately 2 800 students were required for an adequate sample for the written component and approximately 1 000 (50 sites with 10 students at each grade level) students for the practical component.

In March, April and May, 1995 work continued on the finalized versions of the "Bulletin d'information à l'intention des écoles" and the "Guide d'information à l'intention des parents et des élèves." All members of the Science Development Team were involved in the administration of the written and practical assessment instruments in their respective jurisdictions. At every administration site across Canada, teachers and students were asked for their comments about the assessment, the instruments and the items. The team, despite being separated by distance, prepared a detailed scoring guide for the written component of the assessment. Codes were used for the possible correct answers, the anticipated wrong answers, ambiguous answers, wrong answers and no answer. This coding system was designed to determine how the question was operating so that, if necessary, logical changes could be made. Because of this coding system, a mass of extant data exists. It could provide information on specific concepts in science that students know, don't know, and those concepts about which they have misconceptions. With the cooperation of the CMEC secretariat, specifically the National Coordinator of the SAIP, local teachers from Toronto were invited to score the written assessment in June, 1996. To score the practical component of the assessment in July, teachers representing all the jurisdictions were invited

to participate. In Alberta and Ontario, the field test instruments were reviewed by experts and underwent a tolerance and understanding analysis or bias review. In Alberta, the field test instruments were edited by a professional editor in preparation for the final item selection. On May 31, 1995, a draft version of the student questionnaire was circulated from the Ontario team member to other Science Development Team members.

At the same time, in March, the second[30] set of standard-setting meetings for the SAIP assessment instruments was being held across Alberta. Approval for this process was provided by the deputy ministers at the 61st ACDME meeting of February 3, 1995. Standards or expectations for student achievement were being set for the SAIP reading and writing assessment administered in May 1994. The participants were given an overview of the SAIP and the reading and writing assessment and the criteria for each of the five levels. This was followed by a group discussion. The questions the groups answered, using consensus, were as follows: What proportion of 13-year-olds should be performing at each of the levels? and What proportion of 16-year-olds should be performing at each of the levels? Most of the standard-setters were secondary language arts teachers, but representatives from post secondary institutions, parent councils, school boards, teachers' associations, and business and community groups were also included. The representatives from jurisdictions who participated as observers were from Alberta, Saskatchewan, Manitoba, Nova Scotia and the Northwest Territories. A summary report dated May 23, 1995 was circulated to the ministries by the CMEC National Coordinator for SAIP and feedback was sent to Alberta. In November, 1995 a similar standard-setting session occurred in Ontario. The reading and writing standard-setting occurred after the public and technical reports were released. The standard-setting sessions for the science assessment were planned in fall of 1996 so that the public report due for December 1996 release could include both the achievement and the standard-setting results.

---

[30]Alberta had held an earlier set of standard setting meetings (either in 1993 or 1994) to establish standards for the mathematics assessment.

On May 17, 1995, a formal science statistical design meeting was held in Toronto. Members of the Report Development Group, the leader and other members of the Science Development Team, the report writer, the statistician who was a representative from the Ministère de l Éducation, Québec, and the SAIP National Coordinator were in attendance. In April, 1995, preliminary discussions had been held with the statistician, the report writer, the leader of the Science Development Team, and the National Coordinator of the SAIP, in Québec City. These were extremely important meetings as the design of the science assessment was discussed with the statistician. Because the holistic scoring process used by in the reading and writing assessment provided ambiguous information, the statistician advised against having extended-response questions in the written component of the assessment and against holistic scoring of the practical component of the assessment. Instead the Group suggested a scoring guide with criteria that referred to the level attained by the student. A coding system for this criterion-references scoring guide was suggested at the April meeting. By the May meeting, a coding system for the scoring of field test instruments was developed and provided to the statistician. The supervision of the scoring, coding, and data entry was to be the responsibility of the Science Development Team. The statistical analysis which would include classical item analysis and data-cross referencing was to be the responsibility of the Ministère de l Éducation, Québec. With the assistance of Québec, the Science Development Team was responsible for the interpretation of the resulting statistics. Also, plans were made for the full-scale administration targeted for May 1996. As well as the statistical item analysis, Québec was responsible for determining and communicating the sampling procedures to ensure a random sample of students was chosen. The Science Development Team was responsible for providing the coding system for the scoring, for scoring, and for coding. Ontario was responsible for data entry and file preparation. Problems with the data processing in the mathematics and reading and writing assessments were recognized and the decision was made to improve these procedures for the science assessment.

Scoring the Official Field Test

A Science Development Team meeting and the scoring of the short-answer responses of the written component of the field test was held in Toronto, Ontario from June 5 through June 16, 1995. In the first week, all eight members of the Science Development Team worked together to choose exemplars for scoring the practical tasks and to review their responsibilities for the upcoming September 1995 meeting and for the overall success of the assessment. In the second week, six members of the Science Development Team continued the preparations for the scoring of the practical component and two members of the Science Development Team guided and supervised 11 English and French speaking teachers in the scoring and coding of student answers for the short-answer responses of the written component of the field test. As part of the supervision of the scoring, random selections of scored papers were rescored to monitor and ensure reliability. Using a simple analysis of their marking, a member of the Science Development Team noted that the markers demonstrated approximately 91% reliability. In preparation for the scoring of the practical component of the field test a detailed answer key was prepared, per practical task, using a coding system similar to the written component. Two exemplars were chosen to illustrate achievement at levels and two student samples were chosen to allow teachers to practice scoring and to use the coded answer key. Particularly ambiguous samples were chosen. One sample was chosen to be used for a reliability review during the scoring. Exemplars and samples were chosen in English and in French ('sense' translations were added directly to each sample) so that both sets of teachers had the same scoring package. This preparation took approximately one half day per task. A plan was made that two different members of the Science Development Team, one French-speaking and one English-speaking, would direct the teachers in the scoring of each of the practical tasks. Review of the task, the answer key, the exemplars, a practice marking, and consensus building to ensure reliability was the suggested strategy. During this meeting, a standard-setting teleconference was held with some members of the Science Development Team, the

National Coordinator of the SAIP, and the provincial contact from Alberta to make plans for this process.

A June 22, 1995 draft document for the plans for standard-setting was prepared by Ontario. It followed a similar methodology and asked similar questions to the process used by the Alberta standard-setting team for the reading and writing assessment. However, it suggested three regional meetings so as to include representation from across Canada. These meetings were scheduled for August and September of 1996 so that their findings could be available by October 15, 1996 and hence be used in writing of the public report.

On June 27, 1995 a meeting was held in Toronto for the SAIP Science Contacts. The purpose of this meeting was to review the science assessment design, evaluate the field test administration, discuss the sampling procedure for the written and the practical components[31], provide the approximate cost for the jurisdiction to oversample with the practical assessment in preparations for the administration of the full-scale science assessment planned for May 1996. Some responses to the evaluation of the field test administration of the practical tasks were collected at that meeting and some were forwarded afterward with task-by-task feedback. Very helpful information was received and found to be useful in the preparation of the final practical tasks. The ministries were also asked to provide a list of all the science courses offered to 13 and 16-year-old students in their jurisdiction so that the student questionnaire could be completed.

The practical task field test scoring was held in Toronto from July 18 through July 28, 1995. Each of the jurisdictions sent at least one teacher representative (or government official) with the exception of the Northwest Territories. A total of 20 participants were involved in that scoring session. All the participants scored one task at a time, with each task taking about half a day to complete. As a measure of reliability, a portion of the papers of each task was scored twice by different teachers. Each marker was provided with an

[31] Each ministry was asked to provide a list of all their schools that have 13-year-old and 16-year-old students registered and the number of each of these age groups in each school.

identification number. The presentations preparing the teachers for the scoring of each of the tasks and the flow of student booklets were the responsibility of Science Development Team members. This session was one of the first apparently relaxed meetings of the Science Development Team.

The Science Contacts indicated their approval of the test design for the science assessment by July 31, 1995. The ACDME and the CMEC were presented an update of the test design of the SAIP science assessment and the progress of the development of its instruments at its 62nd meeting (August 21, 1995) and at its 68th meeting, respectively (September 18 and 19, 1995 in Whitehorse) by the National Coordinator.

Student and teacher comments, booklet-by-booklet, task-by-task, and of the data sheet, were compiled by August 15, 1995 by the Saskatchewan and Nouveau-Brunswick Science Development Team members for consideration at the September 1995 team meeting. These comments, both English and French, were considered in the item selection process for the construction of the final assessment instruments.

Finalizing the Statistical Sampling Design

On August 29 and 30, 1995 a meeting of the Science Development Team leader, the statistician, and the National Coordinator was held to discuss the SAIP Science Statistical Sampling Design. From the agenda, it appeared that the results from the field tests scoring sessions were available and explained. A small scale field test of the final instruments was suggested for the fall of 1995 which also included field testing the questionnaire. To achieve this, a sample size of 1 000 per age level was suggested. Discussions which followed centered on the statistical methodology for the full-scale assessment including the sampling procedure, setting up data files, scoring guides, scoring and coding, data entry, and statistical analysis. With the exception of the sampling procedures and setting up of the data files the statistical procedures for the final assessment were similar to those used for the field test.

The sampling procedure and data files for the written assessment are described as follows. Data bases were generated by age and language. A designated representative from each jurisdiction was asked to obtain a tally of all the schools that had 13 and 16-year-old students enrolled. The names, code number of the school, and number of students in each population were sent to the statistician. Jurisdictions were allowed to exclude special case schools, as long as no more than 2% of the population were excluded. The statistician then selected the schools that would be involved in the study by a weighted sampling selection, so that large schools had a greater chance of being selected. This list of schools was then sent back to each jurisdiction with a procedure that explained how the students could be chosen. Using a systematic draw, the representative from each jurisdiction then chose the students involved in the assessment. This was one way of obtaining the sample: select the schools and select the students. Two other possible procedures were followed in special case jurisdictions. A census of the schools was involved but a sample of their 13 and 16-year-old students was made or a census of the schools and a census of the 13 and 16-year-old students was taken. No replacement of students in the sample was allowed, however make-up sessions for absent students were allowed. The sampling procedure for the national sample of the practical component of the science assessment follows. Within the sample of schools selected for the written component of the assessment, the statistician chose a sub-sample of schools for the practical component. A school could be selected for both the written and for the practical components of the assessment but small schools were precluded from being selected for both. The jurisdictions, then, selected the students by the systematic selection procedure described above. Since no make-up sessions were available for the practical component, extra students were selected and invited to attend on the day of the assessment. At that time only the required number, generally 10 students wrote the practical assessment, extra students were sent back to their regular classes.

Final Assessment Instruments and Documents

The hectic pace returned at the September 11 through 28, 1995, Science Development Team meeting in Fredericton. With eight members in attendance and guest attendance of the statistician and the National Coordinator (for two of the 15 days) the comments and the statistical information were studied and the final forms of the instruments were made. The National Coordinator reported the CMEC approval, in principle, of the SAIP science criteria, framework, and test design at their 68th meeting.

The test design is described as follows. Test "A" had 12 level three questions, ten multiple-choice and two short-answer questions; with two questions from each of the six domains assessed. (This design was the same as the design for Field Test A, see Table 4-1.) Test "B" had 66 questions, four questions at levels one and two for the four knowledge domains; except level three which had two questions for each domain, and five questions at levels one and two for the nature of science, and the science, technology and society domains; except level three which had three questions for each of those domains. (This design was the same as the design for Field Test B, see Table 4-2.) Test "C" also had 66 questions, four questions at levels four and five for the four knowledge domains; except level three which had two questions for each domain, and five questions at levels four and five for the nature of science, and the science, technology and society domains; except level three which had three questions for each of those domains. (This design was the same as the design for Field Test C, see Table 4-3) Test "B" had 24 short answer and 42 multiple choice questions. Test "C" had 25 short answer and 41 multiple choice questions. The student wrote 26 questions at each of the five levels. A student achieved the level by getting 15 acceptable answers out of the 26. See Appendices G1, G2, and G3 for the blueprints for the "A", "B" and "C" tests. Each question is categorized according to the domain and skill criteria to which it is referenced and is referenced to the number it represents on each test.

Each test took almost two days to construct. The performance of English and French students, male and female students, the overall difficulty of the item, the correlation of the item, item-response theory statistics, the criteria, and teacher and student comments were all considered in the final decisions. One Science Development Team member summarized the construction of the final instruments as follows:

> Right, this was complex. The items had to fit domains, our intuition . . . but we also had to have a statistical basis for justification because the "A" Form is critical here, and that's where the IRT, item response theory, was used to select the form "A" questions. We then went relying on our ability to juggle scenarios and use conventional statistics, item analysis statistics to select the "B" and "C" form questions, but I think given the high stakes nature of the "A" in this low stake's assessment. Like, it's the cutting instrument . . . we wanted to have good questions there. So I think we've pretty well relied on those IRT statistics and on our intuition so the items fit all criteria. They had to be dependable.

The statistics obtained from scoring the 13 practical tasks were studied and by the end of a long day, seven practical tasks were chosen. Modifications based on statistics, teacher comments, and marker suggestions were made on each practical task by different individuals of the Science Development Team, independently, and when completed a dated, hard copy circulated to each member and hard copy and disc was supplied to the Performance Skills Coordinator from Saskatchewan.

Each task had five questions; one question at each level of assessment. A student achieved a skills level by getting four acceptable answers out of the seven. The October 1995 mini-field tests were planned with only the provinces of the Science Development Team participating. These field tests were done to test the administration procedures, the final instruments, the data gathering and processing procedures. This was done to check for unanticipated problems. All accessory documents, such as questionnaire, framework and criteria, information bulletin and brochure, and administration manuals were discussed

and finalized. Likewise, the cover, the data booklet and the instruction page were discussed and finalized. This meeting ended with a group photograph session. One member reflected on this complex and taxing meeting: "Actually, that was one of our best meetings for collaboration not only production, I think everybody said what they had to say and I think we were heard." See Appendix G4 for the design of the practical component of the science assessment.

In each of the four provinces of the Science Development Team, approximately 30 grade nine and 30 grade 12 students wrote the written component and approximately 10 grade nine and 10 grade 12 students wrote the practical component of the nearly final form of the science assessment, in October 1995. All the completed student booklets were sent to Ontario. One of the Ontario members of the Science Development Team scored and coded these student booklets. These data were entered in the computer program and a statistical analysis was run.

In November, all finalized documents were sent to an English editor in Alberta and a French editor in Nouveau-Brunswick. At the end of November, all documents pertaining to the final assessment were mailed to each of the Science Contacts for final analysis and for ultimate approval. Although in previous assessments final approval was required by the ACDME and the CMEC, the revised management structure allowed each Science Contact to represent his or her jurisdiction. Most responses were received from the jurisdictions by the first week of January 1996. Each jurisdiction received a reply from the leader of the Science Development Team itemizing the teams' response to the suggestions on February 6, 1996. All of the concerns raised were addressed either by resulting in changes to the documents or through explanation.

Seven members of the Science Development Team attended the December 4 through 8, 1995 meeting in Regina, Saskatchewan. The assessment instruments were reviewed in light of the field test results. Individual item statistics were analyzed and the items were refined, or replaced with an item from the previous field test. Over the next three days, the

assessment instruments were finalized to the agreement of all team members in attendance. The duties of the Science Contacts, school coordinators, and practical task administrators were refined and the administration manuals reviewed and finalized. Plans were made for the incorporation of comments from the Science Contacts, for printing of all documents, for the upcoming scoring sessions, and for gathering the information required for the technical report of the assessment. By early January all accessory documents to the SAIP science assessment were ready for printing and by the end of February, all test documents were ready for printing.

The Sampling Procedures in Action

On November 2, 1995, a sampling procedures document was sent to the Science Contacts and on November 7, 1995, the Science Contacts were asked to declare inclusions or exclusions in the oversample for the practical component and to return these to Ontario by December 15, 1995. At that same time, they were asked to count the number of science assignment kits they received to allow them to generate the participation data for each form of the assessment. In November, the jurisdictions helped create the Canadian data base of schools with 13-year-old and 16-year-old students.

In December 1995, schools were randomly selected by the statistician and this information was returned to the respective jurisdictions. The schools were contacted and asked for names of all their 13 and 16-year-old students as described in the section entitled "Statistical Sampling Design." Using the sampling procedure provided to each jurisdiction, a random selection of students was made by about the end of March, 1996. The Canadian sample of schools for the practical test was also made by the statistician, and the students for the practical assessment were similarly chosen by the jurisdictions. The participation data[32] were prepared by the jurisdictions and sent to the statistician by May 31,1996.

---

[32]Each representative from the jurisdictions received a software program that listed the names of his/her schools involved. The representative entered the number of students from each school that actually participated in the assessment.

Mail outs for the Full-Scale Assessment

The Science Contacts at each of the jurisdictions received the materials for the full-scale final assessment in a series of five mail-outs that began in January 1996 and finished in March 1996. "The Handbook for Provincial/Territorial Coordinators", written by the Science Development Team member from Ontario was translated, printed, and shipped by the CMEC secretariat in January, 1996. The student sampling procedures prepared by the statistician from Québec were also included with this shipping. The finalized versions of the "Science Assessment Framework and Criteria" and "Bulletin d'information à l'intention des écoles" were translated, printed and shipped by the members of the Science Development Team from Nouveau-Brunswick. The "Guide d'information à l'intention des parents et des élèves" brochures were finalized, translated, and 30 000 English and 15 000 French printed, and shipped, also, by the members of the Science Development Team from Nouveau-Brunswick. This shipping occurred in January. The "Handbook for School Coordinators" was finalized by the Science Development Team Leader in Alberta. This document was then translated, printed[33] and shipped by the CMEC secretariat by February 15, 1996. Also included in this shipping was the "Handbook for Practical Task Coordinators" that was finalized by the Performance Skills Coordinator of the Science Development Team from Saskatchewan which was translated, printed[34] and shipped by the CMEC secretariat. The English and French tests "A", "B", and "C", the "Practical Tasks" test and the "Student Questionnaire" for students were finalized, by the leader of the Science Development Team in Alberta. These were printed and shipped by a contacted computer systems company. The scannable answer sheets for test "A" and questionnaire sheets were designed and printed, and the tracking barcodes were produced in triplicate and also shipped, by this contractor by March 15, 1996. The materials for the kits for the

[33] 2 500 English and 1 100 French copies were printed. This represented sufficient numbers for each school participating in the written component.

[34] 200 English and 200 French copies were printed. This represented sufficient numbers for each school participating in the practical component.

practical tasks were chosen, ordered, assembled, packaged and shipped under the direction of the Practical Task Coordinator in Saskatchewan.

Although not all of the jurisdictions had indicated inclusion or exclusion in an oversample for the practical component, the ordering process began Saturday, December 9 1995, immediately after the December Science Development Team meeting. By December 10th 1995, all requirements were in place and requisitions were submitted to the Saskatchewan government purchasing agency. The assumption was made that three provinces would participate in the oversample; Nouveau-Brunswick - francophone, Nova Scotia, and Saskatchewan. However, on January 15, 1996, Ontario made the decision to oversample. Accordingly, all orders were increased.

As the materials began to arrive, a staff of three began to assemble the 140 kits required for the Canadian sample and the four-province oversample. Each kit contained 17 stations. A kit containing 17 stations of the seven practical tasks was designed so that the ten students could work through all tasks with little or no wait time. The assembly and packaging took three people one month of six long-day weeks. By March 7th, 1996, all 140 kits had been shipped to the jurisdictions. Of all the 140 kits assembled, only three kits had missing items; two sent to Ontario and one sent to Manitoba. The support of the Assistant Deputy Minister of Saskatchewan Education and the relatively smooth bureaucratic process of the government of Saskatchewan facilitated this entire process.

Oversampling in Saskatchewan

The oversample in Saskatchewan involved 100 randomly selected schools with 20 randomly selected students per school. The administration of this oversample required ten practical task administrators. Since Saskatchewan had done performance testing as part of the assessment of their junior high math, health, and elementary science programs, they had expertise in administering performance testing, as it is called in Saskatchewan. In addition, a full day training session was provided to the practical tasks administrators. This

involved taking a look at the framework and criteria, sharing the goals of the assessment, and experiencing a hands-on test of each of the tasks. Since the official decision to oversample in Saskatchewan was not made until February 15th, 1996, two additional people were hired to telephone all the schools and to indicate that they had been selected to participate in this oversample. The names of these schools were obtained from the statistician in Québec from the information provided to him by the Science Contact.[35] The selected schools then sent the names of their 13 and or 16-year-old students to the Saskatchewan ministry, which then used the sampling procedures provided by the statistician from Québec to select the students. This was followed by an administration procedure similar to the one used for the Canadian sample as described in the section entitled "Statistical Sampling Design."

Administration of the Science Assessment

The Science Contacts were responsible for distributing all necessary materials to the school coordinators for the written component of the science assessment. The school coordinator notified the selected students and disseminated the bulletin and the brochure to the participating parents and students, respectively. The Questionnaire and test "A" were also administered to the randomly selected students by the school coordinator. Immediately after the students were finished, the school coordinator marked the test "A" with the scoring guide provided and gave students the appropriate test ("B" or "C") to allow them to continue. Students achieving a score of less than seven out of 12, wrote test B, whereas students achieving a score of greater than six out of 12, wrote test C. The school coordinators calculated the participation rate for the written assessment for their school and then sent all of the materials back to their ministry.

[35]The Science Contact in Saskatchewan always operated on the assumption that Saskatchewan was participating in this until a definitive 'no' arrived from their Assistant Deputy Minister.

The Science Contacts were responsible for selecting a sufficient number of practical task administrators, training them, and supplying them with the names of students, locations of the schools, and the kits for the practical tasks. The practical task administrators made arrangements to enter the school, set up the tasks, supervise the randomly selected students who wrote the practical component, pack up the kits, calculate the participation rate, and return all materials to the ministry. These participating students and their parents also received the Brochure and Bulletin, respectively.

The Science Contacts assembled all the materials from their school coordinators. They entered the participation data for each component of the assessment for their jurisdiction and forwarded this information to the statistician. The completed student questionnaires and tests "A", "B", and "C" were sent to the contracted computer systems company by June 1st 1996.

In the spring of 1996, approximately 1 300 schools with 15 000 13-year-old students and 1 100 schools with 15 000 16-year-old students were involved in the written component of the assessment. Including the four provinces that over sampled, approximately 400 schools with 6 500 13-year-old students and 340 schools with 6 100 16-year-old students were involved in the practical component of the assessment.

SAIP Science Assessment Achievement Standards

On May 2 to 3, 1996, the CMEC secretariat hosted a meeting of the committee for setting national expectations for students in science. Nineteen participants were involved with at least one representative from each jurisdiction. It was chaired by the National Coordinator of the SAIP. The purpose of the meeting was to draft a proposal for establishing national standards for science. The committee decisions included having four regional consultations, involving approximately 100 (total) parents, educators, business, community, and government members. Each consultation was planned to follow the same strategy. Members of the Science Development Team would present the framework and

criteria context for the assessment, the exemplars used for standard setting would be shared, questions would be posed of participants asking them what they believe the proportion of 13 and 16-year-olds should achieve at "x" level, small group discussions would take place and judgments of the participants would be called for. The compilation of the judgments of the approximately 100 participants of the four regional meetings would establish the final achievement standards. This strategy follows the standard-setting procedure used in March 1995 when Alberta hosted a series of standard-setting meetings to determine the standards for the reading and writing assessment. The expectations or the achievement standards set by these participants will be reported with the results of the science assessment. A discussion of the expectations and results will also be included.

Preparation for scoring of the written and practical components

In preparation for the scoring of the written and practical components, a meeting of the entire eight member Science Development Team was held in Toronto, Ontario from June 3 until June 7, 1996. The statistician and the National Coordinator attended the meeting for one day. On this day, problems with incomplete questionnaires, incomplete information provided by the science contact or provincial coordinators, missing assessment components[36] , and the possibility of an item performing poorly were raised but not resolved. The Science Development Team provided guidelines for the statistician but recommended that decisions be made as the severity of the problems become known. One day was spent simplifying the coding system and finalizing the scoring guides for the short-answer questions in the "B" and "C" tests. The strategy for scoring the written component was developed but it was based primarily on the method used in scoring the written component of the field test one year earlier. Newly graduated, local teachers from

[36]Sometimes components of the assessment were misplaced or mistakes in administration were made. For example a student did not write test A but completed a Questionnaire and both test B and C, or other oddities.

Toronto were employed from June 17 to June 28, 1996 to score the short answer questions of the written component.

Next, the team chose four English and four French samples of student work per practical task which would be used to prepare a scoring guide and booklet for the teachers involved in scoring of the practical component. Two samples, one in English and one in French, were chosen to be the exemplars of acceptable student answers for each of the questions at each of the five levels. Two samples, one in English and one in French, were chosen as suitable training or practice responses. These were chosen to allow consensus building discussion to occur so that the scoring guide would be applied the same way by all teachers. Four samples, two in English and two in French, were chosen to be used in a morning and an afternoon reliability review. A division of labour approach was used. Pairs of team members worked on choosing a range of samples for one or two tasks. Then, as a group, final choices were made. This ensured a more efficient process. The strategy for scoring the practical tasks was developed but it was based primarily on the method used in scoring the practical component of the field test one year earlier. One hundred and twenty teachers were involved in scoring the practical tasks in Moncton from July 22 until August 3, 1996.

The responsibilities for running the scoring session of the practical tasks were divided up among the team members. Two would make the training presentations, two would supervise the scanning process, two would organize and ensure the flow of papers, and the remaining two would act as head markers.

Scoring the Written Assessment

Scoring and coding of the short-answer questions of the written component occurred from June 17 through 28, 1996 in Toronto. Approximately 30 000 "B" and "C" tests were scored using the scoring guide and coding system prepared over the last year and finalized in Toronto during the June 3 through 7, 1996 meeting. Forty-eight English and 23

French-speaking newly graduated teachers from Toronto and Toronto-area universities participated in this scoring. For each question scored a letter code was transcribed next to its question number. The possible letter codes A, B, C, and D represented known correct answers. The letter code E represented an unidentified wrong answer. The letter code K represented a wrong answer and the letter N represented no student response. These letter codes for the short-answer questions and the letter codes, A, B, C, D or N for the student answers of the multiple choice questions were recorded in pencil to the left of the number of the question. These prepared "B" and "C" tests were sent to the contacted computer systems company. Data processors were expected to key punch the appropriate code into a prepared software program. As of August 1, 1996 this activity had not yet been completed.

Scoring the Practical Assessment

On December 15, 1995, la directrice, Direction des services pédagogiques, du Nouveau-Brunswick, submitted a proposal to the National Coordinator of the SAIP, to host the marking of the practical tasks in Moncton, Nouveau-Brunswick. Accommodations for the scoring and the markers, and the availability of English and French markers were outlined. This proposal was presented to the ACDME at their 63rd meeting and at the 69th meeting of the CMEC and was accepted. Therefore the scoring of the practical tasks was held in Moncton, Nouveau-Brunswick as planned from July 22 to August 2, 1996.

To complete the marking of the approximately 13 000 papers; 8000 English and 5000 French, 42 French-speaking markers, 66 English-speaking markers and 12 bilingual markers were involved. Of these 120 markers, 20 were head markers. The head markers met one day prior to the start of marking, July 22, 1996. On this day the team and the head markers met to review the scoring guides task-by-task. By the end of that day the head markers were prepared for the scoring of the first four practical tasks. The preparation took longer than expected. On the first day of marking the markers were welcomed by the

Minister of Education, Nouveau-Brunswick, the National Coordinator of the SAIP, the leader of the Science Development Team and the Nouveau-Brunswick team member.

Generally, the scoring of each practical task followed the same pattern. Two team members provided a review of the scoring guide, a look at two exemplars and the opportunity to practice on four training papers. These instructions were provided in both French and English to the entire group. The head markers made clarifications as necessary and handled the questions at their respective tables. Each head marker was responsible for five markers. The papers arrived at the markers' tables in three waves. Of the papers in the first wave, each bundle of approximately 15 papers had five papers that were invisibly marked with ultraviolet ink. After this wave of papers was scored, the head markers led their group through a reliability review. The reliability review consisted of each marker scoring two papers, paper one and paper two. All of paper ones were the same for every marker, as were all of paper twos. For 100% reliability, every marker had to score every question exactly the same. After the reliability review, the markers scored the second wave of papers. These were packaged in bundles of just less than 30. While the markers were scoring this wave of papers, the team members responsible for the paper flow prepared the final wave of papers. The papers that were marked with ultraviolet ink were removed from the first wave of papers. Their already coded scannable answer sheet was removed and replaced with a second uncoded (for that particular task) scannable answer sheet. These marked papers were tucked into the final wave of unscored papers were then scored twice. The results of this double scoring provided a truer indication of how reliable the scoring actually was.

By the end of the first week, three practical tasks were completely scored and a part of each was rescored, blind. In preparation for scoring the remaining practical tasks, the head markers met on Friday, after scoring, and prepared the scoring guide for one more practical task. Intense, and at times, heated discussion occurred, especially about the requirements of level four and five responses. Although these discussions were arduous,

they were important so that the team presenters and the head markers had a clear conception of the scoring guide at each level. This clarity helped set the stage for reliable scoring. In the five days remaining the head markers prepared the scoring guides for the last two practical tasks and the markers scored the final four practical tasks.

Unlike the scoring of the short-answer questions of the written component, the markers were responsible for scoring and coding the score obtained for each question[37] of each practical task on a machine scannable sheet. After all seven of the tasks were scored in each booklet the machine scannable sheet was removed and scanned. The data from this scanning was used to link to the data provided by the student questionnaire, hence the age, jurisdiction, and student characteristics could be identified.

Report writing

The scanning and data processing in preparation for the report writing took place in August 1996. The data processing involved six steps. First the items were analyzed using classical item analysis and item response theory. Second, items were analyzed using differential item functioning to test for bias. Third, statistics such as the means were used to analyze the overall assessment and then each domain. Fourth, the questionnaire was linked to the assessment and cross-tabulations were used in order to look at performance and contextual variables such as gender, language, and population. Fifth, with the linkage to the questionnaire, factor analysis and stepwise regression statistics were used. And, finally sixth, items were grouped with the same level of difficulty and according to each of the five levels of achievement in a process called scaling.

Prior to writing the report the student samples for the practical component were weighted. Since only approximately 10 students were selected from any school that was selected in the sample, the student samples were weighted to represent their population.

---

[37]For each practical task, the student had five questions to answer. Each question tested one of five of the skill levels. Level one questions were low level primary school whereas level five questions were high level secondary or even pre-university questions.

Factored into the weighting formula was the size of the school, the number of students selected for the assessment, and the number of students actually participating in the assessment. The standard-setting was expected to be completed by October 15, 1996. All of this information would be used by the report writers and the Report Development Group to prepare the public and technical report. The public report is expected to be available by the first week of December 1996 for a middle of December 1996 press conference. The technical report is expected to be available to the jurisdictions in the spring of 1997.

ANALYSIS OF THE FINDINGS

Based on reviews of the documentary data, rereadings of the interview transcripts, and much reflection, five themes emerged in the analysis of the genesis, establishment, and the implementation of the SAIP. They reflect what the researcher interpreted as major tensions or issues which influenced and shaped both the context and the process of the SAIP. These were (a) pressures for economic advancement and quality, (b) the strategies of leaders, (c) the value of consensus, (d) the pressures of time and timing, and (e) the presence of change in the program. Each is discussed in turn.

Themes

(a) pressures for economic advancement and quality

The timing of the establishment of the SAIP can be related to a number of factors. The increasing awareness among Canadian deputy Ministers of Education of similar initiatives undertaken by other education systems in the world and Canadian involvement in the CERI project raised interest in the general question of the quality of education in Canada. According to news articles in the Globe and Mail on education, during 1988 and 1989, the general public was more concerned about religion in schools, native and francophone education, violence and racial tensions in the school, and parity for women in education than in wanting to know how their school systems were doing. However, from

1990 on, business and government began stating their concern about the performance of students in the work place and hence the economic impact of the education system.

Questions about and prescriptions for quality, especially as it related to performance indicators and indices were evident in the then current American business literature and discussed in the media. The findings of international assessments which rated students in specific disciplines were used by the media to question the quality of Canadian education. Federal initiatives such as the Economic Council of Canada's report entitled A Lot to Learn (1992) and the Prosperity Initiative (1992) challenged education to provide students who would be better able to help Canada increase its productivity, who would be competent in the generic skills identified by the Canadian Labour Force Training Board (1994) and who would have increased abilities in science and technology

Under Prime Minister Margaret Thatcher mandatory national testing for ages seven, 11, 14, and 16 was introduced in Britain as part of a national curriculum initiative. Similarly, Australia had also begun national discussions about quality schools, a national curriculum, and national performance indicators. Hence, among policy-makers across the developed nations, the trend was toward national standardized testing.

Alberta was a major leader in the institution of the SAIP. This interest could be traced to the philosophy of Alberta Education. It had been putting various management systems in place which would monitor provincial resources and productivity in education. As early as 1982, Alberta Education began focusing on outcomes of education, it instituted exit examinations in key subjects in 1984, and it developed its own indicators system in 198?.

As well, the federal attempts to direct educational curricula and to blame the economic downturn of the eighties on inadequacies in the education system, spurred the CMEC to undertake a program which would simultaneously keep the direction for education in their own hands while obtaining information that could not only answer the worst of the gibes at Canadian schooling but also provide information that would be useful

to educators and parents. Also, between 1985 and 1995, almost every jurisdiction in Canadian had a provincial commission or report on education.

All of these factors not only provided impetus for the development of the SAIP, they also helped keep the process active and on time. However the political milieu also worked to shape how the SAIP was designed and how the eventual data were reported to the public.

(b) the strategies of leaders,

The CMEC's decision to make an education indicators program a priority and part of its action plan was partly based on a proposal written by Alberta Education and presented by the ACDME Subcommittee on Priorities and Programs - ESE. Global economic pressures and concern for the state of and control of education set the stage for the acceptance of such a program. The fact that the proposal was written by an experienced bureaucrat who knew how to use the proper channels to expedite the initiative helped its acceptance. At that time, the chair of the ACDME Subcommittee on Priorities and Programs - ESE was the Deputy Minister of Education, Alberta. He had many years of experience working with the CMEC, first as the Deputy Minister of Advanced Education in the late sixties, then as the Deputy Minister of Education from 1982 until his retirement in 1995. Members of his department were well experienced with the construction of achievement and exit examinations. He played an important role in the CMEC decision to build specific assessment instruments using development teams made up of representatives from interested provinces. On the basis of expertise, the willingness to do the work, and financial backing of a portion of the development, Alberta emerged as the leader of the development teams for the mathematics and the reading and writing assessments, with assistance from Québec.

Likewise, a similar leadership strategy was used by representatives from Alberta to obtain the leadership of the science assessment. The draft proposal outlining the framework for a science assessment, complete with a detailed time line, and budget, was prepared for the fall 1993 CMEC meeting, which awarded the leadership of the science assessment to a representative from Alberta. In both cases, as leader of the Technical Committee on SAIP and of the Development Teams, Alberta provided a conceptual frame of reference that was a product of conscious and coordinated effort. Also in both cases, Alberta was able to hold regular meetings, review the work in progress, bring back 'forgotten ideas', and generally take an array of excellent ideas and change them into a manageable and coherent set. Transforming input into coherence is complicated and difficult. Difficulties arose also because of differences in readiness levels for indicators and indicator systems. But, in both cases Alberta provided a vision and demonstrated the ability to follow through in the face of difficulties. The required support came in the form of mandated approval by the CMEC, an organization that requires consensus and agreement not majority rule.

No other province sought to be the lead province and nor was able to bring that desire to fruition. The original possible leader of science did not bring forward a plausible proposal. As well, the results of the 1990 Request For Proposals seemed to indicate that no other organization had the combination of knowledge, resources, and experience required for its success. It is interesting to speculate that without Alberta's sustained interest, knowledge, and leadership the project might have joined previous ACDME initiatives in much discussion and little action.

(c) the value of consensus

The third theme to emerge was that of commitment to the SAIP by the jurisdictions and Alberta's, CMEC's and the National Coordinator's role in maintaining their commitment. This commitment was gained by using consensus decision making. Although, initially, the establishment of the SAIP was proposed by Alberta Education, the final product of the SAIP was a result of the work of the Technical Committee on SAIP. Also, initially the development teams for mathematics and reading and writing consisted of representatives from Alberta and Québec, but when Ontario wanted to be involved; the development teams were expanded to include their expertise. Since Ontario educates over 50% of the Anglophone students in Canada, their participation was considered politically crucial. Over the year 1991, the CMEC facilitated the creation of the Memorandum of Agreement (Understanding) for the development of the assessment instruments that was agreed to by all jurisdictions. The Memorandum of Agreement outlined a division of labour that maximized expertise and efficiency. The psychometric issues were handed from Alberta to Ontario. The national administration of the mathematics and reading and writing assessments was coordinated by Ontario. According to the Memorandum of Agreement, input from all jurisdictions was requested and obtained throughout the process of instrument development and facilitated by the work of the CMEC secretariat. Consensus decision-making facilitated the participation of the 12 jurisdictions. Decisions made on a majority basis are detrimental to a long-term effort such as the SAIP. Consensus increased the comfort level and commitment to involvement with what was being developed and with the final components. The use of smaller subcommittees or ad hoc groups to bring forward recommendations that were unable to achieve consensus when first discussed was a valuable strategy. The work of the committee was not delayed and yet there was full

discussion on these points. While the committee did not always accept the recommendations, most often this more informal setting and restricted agenda encouraged the development of consensus.

The work of the jurisdictions on the science assessment helps illustrate the variety of decisions and levels of involvement required from all CMEC members. In 1994, a revised Memorandum of Understanding (See Appendix F) for the science assessment was set-up by the CMEC and agreed to by members of the consortium. It also mandated division of labour, input from all jurisdictions and consensus decision making. Essentially the Science Development Team worked jointly with the Science Contacts, representatives from each jurisdictions, over a period of two years to produce science assessment instruments that were satisfactory to all jurisdictions. Besides written input, periodic meetings and an in-service on the practical component of the science assessment allowed a feeling of confidence and acceptance of the products of the Science Development Team to emerge. Representatives of jurisdictions other than the consortium became willing participants in the SAIP science assessment. As a result, students from every jurisdiction participated in the SAIP science assessment. The CERI of the OECD also used consensus decision making strategies to gain commitment and involvement in its international Indicators of Education Systems project.

(d) the pressures of time and timing,

From the fall of 1988, when education indicators became a priority for the CMEC until the fall of 1996 when the report writing of the SAIP science assessment, the last of the discipline achievement assessments, will have taken place, eight year have elapsed. It can be interpreted to be a lot of time to accomplish a little. It can also be interpreted to be a great accomplishment in a short time; the involvement and the commitment of 12 jurisdictions to

the SAIP and the administration of three different assessments, mathematics, reading and writing, and science. A member of the Science Development Team described this aspect of time as follows:

> I don't know. I think on the whole that . . . well we started in March of '94 and it's now June of '96. Over two years you should be able to do an awful lot, but it seems like everything has been so rushed through the whole process. But we got a lot accomplished because the whole first part of the thing just amazed me in how much consultation there had been made with . . . I mean and justifiably, for consensus building in the provinces and territories. Yes, and we're going into their schools and they have control over their schools, so . . . yes they have to be satisfied with what we're doing, but you just never, thinking from outside, you never anticipate how much time that's going to take to get the reactions back, respond to the their comments . . .Get back to them and say, okay, here's what we decided to do. Is that satisfactory? And so on like that. And so, you'd think, well it's two years and four months, or twenty-eight months. We should be able to do an awful lot. Well... I guess we did an awful lot because like we went through two informal and one formal field tests of all these materials. So I think they're pretty, pretty good. And I know that the reaction of teachers in Saskatchewan is really positive to both the written and the practical component . . .

As well, this ambiguous sense of accomplishment can likely be applied to the early establishment of the SAIP.

Three other functions affected the amount of time that elapsed in the genesis, establishment and the implementation of the SAIP. These were the official approval processes of the CMEC, the structured meeting dates of the ACDME and the CMEC and the operation of all CMEC committees, subcommittees, teams, and groups in both official languages. As the Technical Committee reached agreement on aspects of the SAIP, these were brought to the ACDME for preliminary approval. Those aspects that were approved by the ACDME were sent for final approval to the CMEC, while the others were sent back to the Technical Committee to be reworked. This took time. Also, the ACDME and the CMEC generally meet only two times a year, January and February and August and September, respectively. The decision to build assessment instruments specific to the SAIP was an example of a time consuming decision partially due to approval process and meeting dates. Overlaying these limitations was the operation of all meetings and the distribution of all documents in English and French. Questions written in English centres had to be translated into French and questions written in French centres had to be translated into English. Documents written by the francophone members from Ontario and Nouveau-Brunswick had to be translated into English and documents written by anglophone members from Alberta, Saskatchewan, and Ontario had to be translated into French. Feedback and input from the Science Contacts was received in English and in French. Changes to French documents had to be made in the English ones, as well. When the Science Development Team was finished its work two complete sets of document and instruments resulted; one set in English and the other set in French.

The pressures of time came not only from the complex procedures necessary to ensure that all jurisdictions were involved in every aspect of decisions, but also from the

continuing external reports on education and the public demands for standards. If the CMEC did not meet its publicized commitment to the indicators program, it would be taken as further evidence that the Ministers did not really know how their education systems were faring.

(e) the presence of change in the program.

As the program developed over time, constant change was apparent. Neither the final form of the SAIP nor the final form of the science assessment is the same as it was in its initial proposal. Change is most apparent in the management structure of the SAIP. See (Figures 4-2, 4-3, and 4-4 in Chapter IV.) Prior to March 1994, the coordination of the SAIP was carried out by existing members of the CMEC secretariat. In March 1994, a position specific to the SAIP, the National Coordinator for the SAIP, was added to the CMEC secretariat. This person increased the coherence between the functioning of the committees, teams and groups working on the SAIP. Prior to this coordination independent negotiations occurred between them and the ACDME. Subsequently, only items requiring specific approval were taken to the ACDME and onto the CMEC. As a result, decisions that needed to be made by experts were so made and decisions that need to be made by politicians were also so made.

This change in management structure enhanced the quality of the assessment instruments developed by the Science Development Team. Because of the efforts of the National Coordinator, a meeting of the statistician of the Report Development Group and the leader of the Science Development Team was held early in 1995. This meeting was early enough in the development process to allow significant statistical input. Changes to the initial design of the science assessment were made that strengthened the overall

assessment. Similar meetings with the first mathematics, and reading and writing Development Teams were not held. The statistician was not involved until the selection and marking processes. As with the science assessment, in the second cycle of the mathematics assessment meetings have already been held with the statistician and changes to the administration process are already being considered. Also, as the second cycle of the mathematics assessment began the management structure was reorganized again. The researcher found no evidence of meetings held by the SAIP Policy Advisory Committee and the SAIP Administration Management Team, so not surprisingly these committees have either been disbanded or restructured (CMEC, 1996).

The willingness to change structures to expedite the SAIP development of the assessment instruments and, in particular, the appointment of a National Coordinator, would seem to indicate a general desire to ensure that the SAIP was successful. It is also probable that having worked through two assessment development processes, the ACDME was very willing to relinquish the coordination function to one individual. The national education agenda was now the issue that engaged their interest.

That in eight years SAIP moved from being a vague possibility to actual implementation; that it was successful and is on-going and that a federation as large and diverse as Canada was able to achieve this in a spirit of cooperative consensus speaks to the desire of the Ministers of Education to ensure that Canada has the best education system for its children.

Chapter V

SUMMARY, IMPLICATIONS, AND REFLECTIONS

Introduction

The establishment of the SAIP and, in particular, the development of the SAIP science assessment were the foci of this study. Specifically, attention was paid to the leadership of the SAIP, the design, construction, and administration of the framework, criteria, and the instruments of the science assessment, and to the issues and the themes that emerged as the SAIP was established, developed, and was implemented. Finally implications of the SAIP for education in Canada and implications of this study were considered.

To do this research, official and unofficial historical documents were read, analyzed, and interpreted. Particular attention was paid to the years 1988 to 1996, inclusive, however the years prior to 1988 provided some information on the genesis of the SAIP. In addition to this document search, nine formal interviews were held with key informants. These were people who had been or have current involvement with the SAIP. Four were from Alberta, one from Saskatchewan, one from the CMEC secretariat, and three from Ontario. Two interviewees from Ontario had experience with the SAIP as former members for British Columbia and Saskatchewan, while the third participant from Ontario had been a member of the CMEC secretariat just prior to the establishment of the SAIP. The interviews were transcribed, analyzed, and interpreted. Follow-up conversations, by telephone, e-mail, and or in-person, helped clarify ambiguous information or missing details. Also, an unrecorded, informal interview with the statistician from Québec helped confirm the researcher's understanding of the statistical procedures used in the SAIP.

A chronological summary of events of CMEC activities as they relate to the SAIP is presented in this chapter. As well, from analysis of these data five themes emerged. They

related to global pressures, leadership, consensus, time, and change. Additionally, implications for the next science assessment, implications of the SAIP on education in Canada and a discussion of the implications of this research are presented at the end of this chapter. This chapter ends with the researcher's personal reflections on this study.

Chronological Summary of Events

1988

The major accomplishment of this year was the establishment of a Canadian education indicators program, with a three phase plan for implementation. This program was to increase accountability and assess performance, quality and effectiveness of the jurisdictions' educational programs in comparison to Canadian standards and to each other. Another accomplishment was the joint effort of the CMEC and Statistics Canada to improve education statistics data with the development of the Deputy Ministers' Steering Committee on Canadian Education Statistics (1987-88), its two Technical Committees (Elementary and Secondary Education and Postsecondary Education) and a protocol pending.

1989

The two of three parts of the CMEC School Achievement Indicators Project were launched in 1989 to determine if Canada was meeting its goals for its education systems. The planning and design of the involvements indicators was completed and approved. The form and content of the assessment instrument remained to be decided, but this indicator was approved in principle. However, the expectations and satisfaction indicators remained under discussion. The Statistics Canada and the CMEC protocol was signed in September 1989 and the first education statistics report was due in 1990; hence a Canadian Education Statistics Program was initiated. The OECD's CERI was working on an international education indicators project to which Canada contributed. Alberta and Québec were the lead provinces in Canada's participation in the CERI's international indicators. In September

1989, Canadian educators offered a workshop on the establishment of an education indicators systems as part of the CERI international indicators conference in Semmering, Austria.

1990

The CMEC's Annual Report described the SAIP as a flagship program, however, the SAIP dominated the activities of the CMEC this year. The name was now the School Achievement Indicators Program not Project, probably an indication of a possible long term acceptance. As part of a communications strategy, a nongovernmental organization meeting was held in Toronto in June 1990, co-chaired by the Director-General of the CMEC and the Minister of Education from Alberta. A coordinator was appointed for the SAIP from the CMEC secretariat in June 1990. Data gathering began this year on the first group of indicators, the involvement indicators: participation, retention, and graduation rates. This was harmonized with the Canadian Education Statistics Council. Its first meeting was in January 1990 and its first report the Statistical Portrait of Elementary and Secondary Education in Canada (1990) was released in 1990. The second group of indicators- curricular expectation, student, teacher and parental expectations and satisfactions- were completely revised and reconsidered. The end result was the decision to build curricular expectations into the assessments and reconsider the others at another time. Also, background data was to be collected from students and teachers participating in the assessment as a result of these decisions. This descriptive information about expectations was planned to be used to help interpret the results of the assessment. After substantial time had been spent determining the best alternative for the assessment instrument, the work began on the literacy and numeracy achievement indicators. In 1990, a Request for Proposals for the construction of the assessment instruments proved futile. As a result, an Alberta and Québec led consortium was established in September 1990 to manage the construction of assessment instruments. Also, a SAIP Policy Advisory Committee was

formed at that time. In December 1990, Ontario withdrew its active support from the program but retained observer status.

1991

In 1991, the CMEC tried to reach consensus on outstanding SAIP issues so as to involve all jurisdictions in the assessments. In late 1991, a Memorandum of Understanding was drafted to deal with Ontario's concerns about sampling and the use and publication of the results. Data were gathered on student involvement for the years 1989-90 and 1990-91 from each of the jurisdictions; the terms were defined and the Canadian Education Statistics Council planned to report SAIP involvement data in its second report; the Statistical Portrait of Elementary and Secondary Education in Canada (1992). A SAIP bulletin was planned for release in November 1991, as part of a communication strategy. Meanwhile, by June 1991, the consortium for developing the SAIP assessment instruments prepared the framework and criteria documents for the reading and writing and mathematics assessments. Item development for the mathematics assessment followed in preparation for field testing in spring of 1992. A Canadian delegation, led by Alberta's Assistant Deputy Minister of Education, attended the second conference on the international indicators project sponsored by OECD's CERI in Lugano, Switzerland. Canada continued its involvement in this international indicators project.

1992

Based on the Memorandum of Understanding signed late in 1991, Ontario officially joined the consortium responsible for developing the assessment instruments. Field tests of the mathematics assessment were administered in the spring of 1992. Scoring of the field tests was done and final instrument preparations were made for the mathematics assessment during the remainder of 1992. Work on the reading and writing assessment proceeded more slowly. Ontario coordinated the sampling procedures and organized for the 1993 administration and the scoring of the final instruments of the mathematics assessment.

Saskatchewan was an observer, not an active participant in the SAIP. The Canadian Education Statistics Council published the 1992, second edition, of A Statistical Portrait of Elementary and Secondary Education in Canada. Canada continued its involvement in the third phase of the international indicators project of the CERI.

1993

The first Canada-wide assessment of school achievement was administered in the spring of 1993. Approximately 50 000 students across all jurisdictions, except Saskatchewan, wrote the SAIP mathematics assessment. The public report of the mathematics assessment was published in December 1993. Also, a field test of items prepared for the reading and writing assessment took place. In September 1993, the SAIP was expanded to include a science assessment planned for 1996 and that fall the work began to establish a consortium to develop the science assessment. In September, the Joint Declaration: Future Direction for the Council of Ministers of Education, Canada set up an action plan for the CMEC, primarily by setting a national education agenda. As a result of this Joint Declaration, a Pan-Canadian Education Indicators Program and a continued partnership with Statistics Canada and the CMEC were proposed by the Canadian Education Statistics Council.

1994

The consortium for the science assessment was established. It included Alberta, Saskatchewan, Ontario, and Nouveau-Brunswick. A Memorandum of Understanding was developed (Appendix F), a consortium agreement established, and development team members chosen. Over this year, the members of the Science Development Team worked with the Science Contacts in a consultative manner and the framework and the criteria of the science assessment were developed. The framework and criteria were presented to a meeting of nongovernmental organizations and experts for additional feedback. During the

summer of 1994 item writing took place in centres all across Canada. In the fall, the products from the item writing sessions were used to construct unofficial field tests. These field tests were only carried out in the jurisdictions of the members of the consortium. From the statistical results of the unofficial field test held in October, the official field tests were put together for the various written components and for the practical component of the SAIP science assessment.

Also in this year, the First National Consultation on Education was held in Montreal in May 1994. This was a direct result of the formation of the national education agenda called for in the Joint Declaration. The second national assessment of school achievement was administered in the spring of 1994. Approximately 58 000 students across all jurisdictions, again except Saskatchewan, wrote the SAIP reading and writing assessment. The public report of the reading and writing assessment was published in December 1994.

1995

This year was a busy year for the Science Development Team. Early in the year, the instruments for the official field tests were finalized. The plans, orders, packaging, and distribution of the materials required for the practical component were completed. An in-service was held for Science Contacts to assist the jurisdictions with the administration of the practical component of the assessment. The official field tests took place in the spring in every jurisdiction in Canada. Scoring of the written component took place in June and scoring of the practical component took place in July. The statistical instrument and sampling design was put in place. Using the statistical design and the statistical results from the scored field tests the final instruments of the science assessment were constructed. They were then circulated to the Science Contacts for input and feedback and subsequent changes were made. Also by December the Science Development Team had prepared most of the associated documents for the administration of the assessment. Plans were begun for the

standard-setting process for the science assessment, primarily, by the provincial coordinator from Alberta.

The consortium responsible for administering the second mathematics assessment held its first planning meeting in the fall of 1995. An action plan for the Pan-Canadian Education Indicators Program was approved in September 1995. The CMEC also approved the start-up of the Pan-Canadian Science Project in September 1995. The main objective was to develop a common science curriculum by April 30 1997. In November 1995 the Report on Education in Canada was published. It included the results from the two SAIP assessments and the data published in A Statistical Portrait of Elementary and Secondary Education in Canada.

1996

The third national assessment of school achievement was administered in the spring of 1996. Approximately 43 000 students across all jurisdictions in Canada wrote the SAIP science assessment. To accomplish this, the members of the Science Development Team prepared and finalized all documents for the administration, the provincial coordinators selected the random sample of students participating, the team members from Saskatchewan prepared the 140 kits necessary for the administration of the practical tasks, and various ministries of the consortium mailed out all necessary materials to the jurisdictions. The administration was then completed by the provincial or territorial coordinator for the SAIP. After the instruments were administered, all components were returned to Ontario. The written component of the assessment was scored in Toronto in June and the practical component was scored in Moncton in July. The statistical analysis was completed in August and plans were underway for standard-setting in September so as to provide additional information for the report writers. The public report of the science assessment is planned to be published in December 1996.

Also in 1996, work continued on the second mathematics assessment, the Pan-Canadian Science Project, and the Pan-Canadian Education Indicators Program. The Second National Consultation on Education was held in Edmonton in May of 1996.

Implications of the SAIP

The genesis, establishment, and implementation of the SAIP has had at least one significant effect on the operation of the CMEC and several effects on the output of the CMEC and other similar bodies. One significant effect on the operation of the CMEC is best summarized by the Director General, Dr. Francis R. Whyte (CMEC, 1995, p. 5) in his message in the 1992-93 Annual Report.

> There is no question that 1992-93 was a turning point in Council history, and a significant step forward in terms of the creation of collective national leadership in education. . . . The School Achievement Indicators Program broke new ground in April when it carried out Canada's first-ever national testing in schools by administering the assessment instruments in mathematics to a sample of some 56 000 students across Canada. . . . which constituted the most formidable challenge the Council had ever taken up. Its success, however, stands as tangible proof that the provinces and territories can work together on national programs in a sustained manner.

That same year the CMEC approved the Joint Declaration: Future Directions for the Council of Ministers, Canada. The action plan of this agreement was the national education agenda, an agenda that called for collaboration and cooperation on education across the nation. The Council's success with SAIP had an effect on its activities such as the following. As part of this national agenda two national consultations on education were held, one in Montreal, in May 1994 and the other in Edmonton, May 1996. The SAIP was expanded to include science and a mandate was approved for the second cycle of assessment to occur with the second mathematics assessment for 1997, the second reading and writing assessment for

1998, and the second science assessment for 1999. A Report on Education in Canada was published and released in November 1995. The partnership with Statistics Canada and the CMEC was extended and the Pan-Canadian Education Indicators Program was proposed. CMEC's Canadian Education Statistics Council and Statistics Canada have recently published their third edition of A Statistical Portrait of Elementary and Secondary Education in Canada in April 1996.

The CMEC approved the action plan for the Pan-Canadian Education Indicators Program at their September 1995. Some of the components of this Pan-Canadian program appear very similar to the components of the 1988 proposal for the SAIP. Six groups of indicators of the Pan-Canadian Education Indicators Program include indicators on accessibility, student flows (participation, retention, graduation, diploma attainment), school to work transition, achievement, citizenship and satisfaction with education systems. Also part of the 1993 national education agenda, was the proposal for the Pan-Canadian Protocol for Collaboration on School Curriculum. In 1995 the Protocol was approved and the Pan-Canadian Framework of Goals and Outcomes for Science K-12 was established. Its development requires the same consensus-style decision making based on input from all 12 jurisdictions. The products of this collaboration, due 1997, will include a framework of goals and outcomes for science for grades three, six, nine, and 12 and should complement the work of the SAIP science assessment. The lead jurisdictions are British Columbia and francophone Manitoba and their operation is similar to the operation of the SAIP Assessment Development Teams.

Although not CMEC initiatives, two interprovincial agreements, the Western Protocol and the Atlantic Provinces Memorandum of Understanding illustrate joint initiatives in curriculum development. Established in the 1990s neither has any direct link to the SAIP however the spirit of cooperation and collaboration of the SAIP is evident in these agreements. Indirectly, then a collaborative project such as the SAIP made cooperation the norm across Canada. The most striking element of the work in establishing the SAIP and

developing the SAIP science assessment was the birth of the 'culture' of large-scale output from the CMEC.

The implications of the institution of an achievement indicators program are more subtle. They can best be summarized by these comments made by the former Deputy Minister of Alberta Education.

> I do know that the provinces that don't do well are certainly looking at their whole instruction program. They're looking at their standards and whether or not they're communicating their standards. None of them want to be at the bottom. I haven't heard anyone say that. Even though its not designed to compare provinces, they do compare. It's just inevitable. So, one of the benefits is that all of Canada, I think, is moving toward a better educational system.

Likely, the jurisdiction representatives working on the Pan-Canadian Framework of Goals and Outcomes for Science K-12 will consider the results of the SAIP science assessment. The report on the science assessment is timed such that it will be released (December, 1996) before the final framework of goals and outcomes for science are due to be completed (1997). The Memorandum of Understanding for the work of the consortium on the Pan-Canadian Framework of Goals and Outcomes for Science K-12 includes on-going communication with the SAIP Science Development Team. A holistic view of science education across jurisdictions, such as is currently occurring, will serve to improve science education in Canada.

Implications of this study

Although history is constructed reality and a historian attempts to discover data instead of creating it, historical research can assist the educator or administrator in predicting future trends or analyzing similar phenomena in the present. (Borg & Gall, 1989) This research study can offer a method of operation for others undertaking large-

scale projects which cross boundaries of authority. Where the participants are "volunteers," the leader must operate with a greater attention to cooperation than in situations where the participants are mandated to be involved.

Also, this research study can provide specific background information to the committee preparing for the upcoming 1999 science assessment. Perhaps, after the Pan-Canadian Science Project is completed in 1997, the new Science Development Team may reanalyze the criteria and shift the perspective of the assessment accordingly. The new Science Development Team should recognize that eight members were part of the original development, each representing one of four jurisdictions and one of three disciplines in science. Each member has made a slightly different contribution to the original science assessment, and therefore would have slightly different recommendations for the next assessment.

This study serves to remind those reading, analyzing, or even embarking on a quantitative study, the qualitative aspects involved. The framework, the criteria, the statistical design, and the items that were the basis of the SAIP science assessment were a result of decisions that were made with the information and knowledge available. These were decisions that were shaped by the feedback and input from the representatives at the various ministries across Canada. With different information and knowledge, different feedback and input, and perhaps, different people in the ministries and on the Science Development Team, the framework, the criteria, and the resulting instruments, likely, would have been different.

This study raises the awareness of those reading the results of international assessments or any large-scale assessment. Questions such as: What was the statistical sampling design? What was the participation rate? Who participated? To what degree did all students in the country participate and in which components of the assessment? - need to be asked for any similar survey. Factors such as these were held constant in the SAIP assessments, however international assessments allow for greater choice, therefore

country-to-country comparisons are difficult to impossible. Even questions like which countries participated must be considered. For example, Japan and Korea, rarely if ever, participate in the same assessment, which makes comparison unlikely.

Reflections

This study, the history of the SAIP, serves as a wonderful memento for the researcher, as she was a member of the SAIP Science Development Team, hence the experience is forever captured. However, the excitement of working with members across the provinces and territories is likely not captured. As a participant observer, the researcher can fairly say that the commitment and drive that was evident in all the participants in whatever aspect of making the SAIP a success is tied to this excitement and sense of adventure. Interprovincial association and cooperation is exciting in and of itself, however equally as exciting is the thrill of being a pioneer. The SAIP was a new program in Canada and the first national assessment in Canada. Being a member of any of the committees or teams in a situation like this increases the commitment to the project and increases the chances for its success.

How the various committees and teams operated is interesting, as well. Although Alberta served as the leader for establishing the SAIP and for developing the instruments, all members representing the provinces and territories were equal. Because education in Canada is a provincial responsibility, by constitution, no one province had any mandated authority over another. The committees and teams worked with their leader under the spirit of cooperation, much like volunteers do, resulting in a national endeavour. Without this spirit, developing a national project would be extremely difficult.

The working conditions of the committees and teams served to develop this spirit of cooperation. The meetings were held at sites across Canada, usually far from home. A large amount of work had to be completed in a short time as the meetings lasted from two days to as long as 15 days. Long days of work gave way to long dinners, of course with

the same people. Getting to know the members from the other jurisdictions, socially, helped facilitate the cooperation required to get agreement, first, then get the work completed. This could only happen in Canada because of its size; going home every night after a meeting is simply not possible.

Canada is a nation that is committed to universal education, with breadth and depth. The SAIP is an achievement indicators system that serves to provide information about how universal education in Canada is performing. The SAIP also provides information for interprovincial comparisons. From both the similarities and the differences in their performances, policy-makers in the provinces and territories can make well-founded decisions about their education systems, and hence Canada's.

References

Alberta Education. (1993). Achieving Quality. Final Report of the Educational Quality Indicators Initiative. Edmonton, AB.

Alberta Education. (1988 December). CMEC School Achievement Indicators Project. Unpublished paper.

Alberta Education. (1989 February). CMEC School Achievement Indicators Project. Unpublished paper.

Alberta Education. (1989a May). CMEC School Achievement Indicators Project. Descriptive Framework. Unpublished paper.

Alberta Education. (1989b May). CMEC School Achievement Indicators Project. Phase I Methodology. Unpublished paper.

Alberta Education. (1989c May). CMEC School Achievement Indicators Project. Phase III Methodology. Unpublished paper.

Alberta Education. (1989 November). CMEC School Achievement Indicators Project. Phase II Methodology. Unpublished paper.

Alberta Education. (1990 May). CMEC School Achievement Indicators Project. Proposed Methodology for Studying Society's Expectations and Satisfaction Related to Schooling in Canada. Unpublished paper.

Alberta Education. (1989 November). CMEC School Achievement Indicators Project. Phase II Methodology. Unpublished paper.

Alberta Education. (1994 February). Criteria for Assessing Scientific Inquiry Skills. Unpublished paper

Alberta Education. (1993 September). Scientific Literacy Assessment. Unpublished paper.

Alberta Education. (no date). TIMSS Canada. [Brochure]. Edmonton, AB.

American Association for the Advancement of Science. (1990). Science for all Americans. New York: Oxford University Press

American Association for the Advancement of Science. (1993). Benchmarks for Science Literacy. New York: Oxford University Press

Berg, B.L. (1995). An Introduction to Content Analysis. In K. Hanson (ed.), Qualitative Research Methods for the Social Sciences (pp. 174-199). Boston, MA: Allyn and Bacon.

Borg, W.R., & Gall, M.D. (1989). Educational Research An Introduction (5th ed.). New York: Longman.

Bosetti, R. (1989, September). Implementation and Use of Education Indicators: The Canadian Experience. Paper presented at the General Assembly of the International Education Indicators Project, CERI, OECD, Semmering, Austria.

Byfield, V. (1988, November 7). Contemplating national testing. Alberta Report, 26.

Canadian Labour Force Development Board. (1994). Putting the Pieces Together: Toward a Coherent Transition System for Canada's Labour Force. Ottawa, ON.

Council of Ministers of Education, Canada. (1988a). A Critical Analysis of Provincially Sponsored Student Achievement and System Evaluation Programs Examined From a Canada-Wide Perspective, Unpublished paper.

Council of Ministers of Education, Canada. (1988b). Agreed Memorandum on a Council of Ministers of Education, Canada. Toronto, ON.

Council of Ministers of Education, Canada. (1994a). Agreed Memorandum on a Council of Ministers of Education, Canada. Toronto, ON.

Council of Ministers of Education, Canada. (1988c). An Overview of Provincially Sponsored Student Achievement and System Evaluation Programs in Canada. Unpublished paper.

Council of Ministers of Education, Canada. (1975). Annual Report 1974 - 1975. Toronto, ON.

Council of Ministers of Education, Canada. (1976). Annual Report 1975 - 1976. Toronto, ON.

Council of Ministers of Education, Canada. (1977). Annual Report 1976 - 1977. Toronto, ON.

Council of Ministers of Education, Canada. (1978). Annual Report 1977 - 1978. Toronto, ON.

Council of Ministers of Education, Canada. (1979). Annual Report 1978 - 1979. Toronto, ON.

Council of Ministers of Education, Canada. (1981). Annual Report 1979 - 1980. Toronto, ON.

Council of Ministers of Education, Canada. (1982). Annual Report 1980 - 1981. Toronto, ON.

Council of Ministers of Education, Canada. (1983). Annual Report 1981 - 1982. Toronto, ON.

Council of Ministers of Education, Canada. (1984). Annual Report 1982 - 1983. Toronto, ON.

Council of Ministers of Education, Canada. (1985). Annual Report 1983 - 1984. Toronto, ON.

Council of Ministers of Education, Canada. (1986). Annual Report 1984 - 1985. Toronto, ON.

Council of Ministers of Education, Canada. (1987). Annual Report 1985 - 1986. Toronto, ON.

Council of Ministers of Education, Canada. (1988d). Annual Report 1986 - 1987. Toronto, ON.

Council of Ministers of Education, Canada. (1989). Annual Report 1987 - 1988. Toronto, ON.

Council of Ministers of Education, Canada. (1990). Annual Report 1988 - 1989. Toronto, ON.

Council of Ministers of Education, Canada. (1991). Annual Report 1989 - 1990. Toronto, ON.

Council of Ministers of Education, Canada. (1992). Annual Report 1990 - 1991. Toronto, ON.

Council of Ministers of Education, Canada. (1993). Annual Report 1991 - 1992. Toronto, ON.

Council of Ministers of Education, Canada. (1995). Annual Report 1992 - 1993. Toronto, ON.

Council of Ministers of Education, Canada. (1996). Biennial Report 1993 - 1995. Toronto, ON.

Council of Ministers of Education, Canada. (1996). Bulletin d'information à l'intention des écoles. Unpublished paper.

Council of Ministers of Education, Canada. (1994 January). CMEC School Achievement Indicators Program Science Assessment. Unpublished paper.

Council of Ministers of Education, Canada. (1994 February). CMEC School Achievement Indicators Program Science Assessment. Unpublished paper.

Council of Ministers of Education, Canada. (1994 March). CMEC SAIP Science Assessment Framework. Unpublished paper.

Council of Ministers of Education, Canada. (1994 May). Council of Ministers of Education, Canada, School Achievement Indicators Program Science Assessment Framework and Criteria. Unpublished paper.

Council of Ministers of Education, Canada. (1994 June). Council of Ministers of Education, Canada, School Achievement Indicators Program Science Assessment Framework and Criteria. Unpublished paper.

Council of Ministers of Education, Canada. (1994 July). Council of Ministers of Education, Canada, School Achievement Indicators Program Science Assessment Framework and Criteria. Unpublished paper.

Council of Ministers of Education, Canada. (1994 August). Council of Ministers of Education, Canada, School Achievement Indicators Program Science Assessment Framework and Criteria. Unpublished paper.

Council of Ministers of Education, Canada. (1994 October). Council of Ministers of Education, Canada, School Achievement Indicators Program Science Assessment Framework and Criteria. Unpublished paper.

Council of Ministers of Education, Canada. (1995 February). Council of Ministers of Education, Canada, School Achievement Indicators Program Science Assessment Framework and Criteria. Unpublished paper.

Council of Ministers of Education, Canada. (1996). Guide d'information à l'intention des parents et des écoles. Unpublished paper.

Council of Ministers of Education, Canada. (1994b). Guidelines for the Construction of Questions. Unpublished paper.

Council of Ministers of Education, Canada. (1996). Handbook for Practical Task Coordinators. Unpublished paper.

Council of Ministers of Education, Canada. (1996). The Handbook for Provincial/Territorial Coordinators. Unpublished paper.

Council of Ministers of Education, Canada. (1996). Handbook for School Coordinators. Unpublished paper.

Council of Ministers of Education, Canada. (1995). Information for Field Test Administrators. Unpublished paper.

Council of Ministers of Education, Canada. (1993). Joint Declaration Future Directions for The Council of Ministers of Education, Canada. Toronto, ON.

Council of Ministers of Education, Canada. (1986a). Liaison Vol. 10. Toronto, ON.

Council of Ministers of Education, Canada. (1986b). Liaison Vol. 10. Toronto, ON.

Council of Ministers of Education, Canada. (1986c). Liaison Vol. 11. Toronto, ON.

Council of Ministers of Education, Canada. (1987a). Liaison Vol. 12. Toronto, ON.

Council of Ministers of Education, Canada. (1987b). Liaison Vol. 12. Toronto, ON.

Council of Ministers of Education, Canada. (1988e). Liaison Vol. 13. Toronto, ON.

Council of Ministers of Education, Canada. (1988f). Liaison Vol. 13. Toronto, ON.

Council of Ministers of Education, Canada. (1989a). Liaison Vol. 14. Toronto, ON.

Council of Ministers of Education, Canada. (1989b). Liaison Vol. 14. Toronto, ON.

Council of Ministers of Education, Canada. (1990a). Liaison Vol. 15. Toronto, ON.

Council of Ministers of Education, Canada. (1990b). Liaison Vol. 15. Toronto, ON.

Council of Ministers of Education, Canada. (1991). Liaison Vol. 16. Toronto, ON.

Council of Ministers of Education, Canada. (1995). Liaison Vol. 17. Toronto, ON.

Council of Ministers of Education, Canada. (1982-1995). Minutes of the 32 d until the 62nd Meetings of the ACDME. Toronto, ON.

Council of Ministers of Education, Canada. (1985-1995). Minutes of the 46th until the 68th Meetings of the Council. Toronto, ON.

Council of Ministers of Education, Canada. (1989). Minutes of the January 17, 1989 and November 9-10, 1989 Meetings (2) of the Technical Committee on Canadian Education Statistics. Toronto, ON.

Council of Ministers of Education, Canada. (1989-1990). Minutes of the June 1, 1989 until June 14-15, 1990 Meetings (6) of the Technical Committee on the CMEC School Achievement Indicators Program. Toronto, ON.

Council of Ministers of Education, Canada. (1989-1990). Minutes of the November 15, 1989 until May 17, 1990 Meetings (3) of the subcommittees of the Technical Committee on the CMEC School Achievement Indicators Program. Toronto, ON.

Council of Ministers of Education, Canada. (1993-1995). Minutes of the December 13-14, 1993 until June 27, 1995 Meetings (5) of the Science Contacts for the CMEC School Achievement Indicators Program science assessment. Toronto, ON.

Council of Ministers of Education, Canada. (1994-1995). Minutes of the December, 1994 until August 29-30, 1995 Meetings (3) of personnel planning the statistical

design of the CMEC School Achievement Indicators Program science assessment. Toronto, ON.

Council of Ministers of Education, Canada. (1995-1996). Minutes of the March, 1995 and May 2-3, 1996 Meetings (2) of personnel planning the standard setting of the CMEC School Achievement Indicators Program science assessment. Toronto, ON.

Council of Ministers of Education, Canada. (1994-1995). Minutes of the March 21-25, 1994 until July 22-August 3, 1996 Meetings (16) of the Science Development Team for the CMEC School Achievement Indicators Program science assessment. Toronto, ON.

Council of Ministers of Education, Canada. (1996). Practical Tasks. Unpublished test.

Council of Ministers of Education, Canada. (1989). Press Release. Education Ministers and Statistics Canada Sign Agreement on Improving Education Statistics. Toronto, ON.

Council of Ministers of Education, Canada. (1995). Press Release. First Report on Education in Canada Released. Toronto, ON.

Council of Ministers of Education, Canada. (1995). Press Release. OECD Releases Third Education Indicators Publication. Toronto, ON.

Council of Ministers of Education, Canada. (1995). Report on Education in Canada. Toronto, ON.

Council of Ministers of Education, Canada. (1993). SAIP Mathematics Assessment 1993. Toronto, ON.

Council of Ministers of Education, Canada. (1994c). SAIP Reading and Writing Assessment 1994. Toronto, ON.

Council of Ministers of Education, Canada. (1994a May). SAIP Science Assessment Criteria. Unpublished paper.

Council of Ministers of Education, Canada. (1994b May). SAIP Science Assessment Criteria. Unpublished paper.

Council of Ministers of Education, Canada. (1996). Science Assessment Framework and Criteria. Unpublished paper.

Council of Ministers of Education, Canada. (1995). Statistical Sampling Design. Unpublished paper.

Council of Ministers of Education, Canada. (1996). Student Questionnaire. Unpublished test.

Council of Ministers of Education, Canada. (1996). Test A. Unpublished test.

Council of Ministers of Education, Canada. (1996). Test B. Unpublished test.

Council of Ministers of Education, Canada. (1996). Test C. Unpublished test.

Council of Ministers of Education, Canada & Statistics Canada. (1989 September). Protocol Between Statistics Canada and the Provincial Governments for the Establishment of a Canadian Education Statistics Program. Toronto, ON.

Council of Ministers of Education, Canada & Statistics Canada. (1990). A Statistical Portrait of Elementary and Secondary Education in Canada. Ottawa and Toronto, ON.

Council of Ministers of Education, Canada & Statistics Canada. (1992). A Statistical Portrait of Elementary and Secondary Education in Canada. Ottawa and Toronto, ON.

Council of Ministers of Education, Canada & Statistics Canada. (1996). A Statistical Portrait of Elementary and Secondary Education in Canada. Ottawa and Toronto, ON.

Dinning, J. (1990, June). The School Achievement Indicators Project. Address presented at the meeting of nongovernmental organizations, Toronto, ON.

Economic Council of Canada. (1992). A Lot to Learn: Education and Training in Canada: A Statement by the Economic Council of Canada. Ottawa, ON.

Hochachka, N. (1989). A Discussion Paper on Literacy. Unpublished paper. Alberta Education.

McEwen, N. (1995a). Accountability in Education in Canada. Canadian Journal of Education, 20, 1, pp. 1-17.

McEwen, N. (1995b). Educational Accountability in Alberta. Canadian Journal of Education, 20, 1, pp. 27-44.

Maisonneuve, Daniel.(1989). Developing Student Flow Indicators Based on Data From One Year. Ministère de l'Éducation du Québec. Québec, CQ.

OECD. (1992). The OECD International Education Indicators, A Framework for Analysis. Paris, France.

OECD. (1994). Making Education Count. Developing and Using International Indicators. Paris, France.

OECD. (1995a). Indicators of Education Systems. Measuring the Quality of Schools. Paris, France.

OECD. (1995b). Indicators of Education Systems. Education at a Glance, OECD Indicators. Paris, France.

Science Council of Canada. 1984. Science for Every Student, Report 36. Hull, CQ

Semple, B.M. (1992). Performance Assessment: An International Experiment. Educational Testing Service. Princeton, NJ.

Shavelson, R.J., Carey, N.B., & Webb, N.M. (1990). Indicators of Science Achievement Options for a Powerful Policy Instrument. Phi Delta Kappan, 71, 9, 692-697.

SOED. (1995). Science Feedback 2. HMSO. Edinburgh, Scotland.

Spradley, J.P., (1979). The Ethnographic Interview. New York: Holt, Rinehart and Winston.

Statistics Canada (1991). School Leavers Survey. Ottawa, ON.

Statistics Canada (1989). Survey of Literacy Skills used in Daily Activities. Ottawa, ON.

Stevenson, H. ((1993, October). Stevenson Study of Mathematics Achievement in Alberta Schools, Summary Report. Based on a presentation by Dr. Harold Stevenson, Edmonton, AB: Alberta Education.

Supply and Services, Canada. (1992). An Overview of Selected Studies on Canada's Prosperity and Competitiveness/Industry. Science and Technology, Canada. Ottawa, ON.

Wiggins, G. (1989). A True Test: Toward More Authentic and Equitable Assessment. Phi Delta Kappan, May, 703-713.

Wiggins, G. (1993). Assessment: Authenticity, Context, and Validity. Phi Delta Kappan, November, 200-214.

APPENDIX A

The

Invitation to Participate

in the

Interview

May 29, 1996

Dear

Further to our conversation in early May, 1996, this is to provide more information about the interview that we spoke of at that time. As discussed, I will be contacting you by telephone on June 9, 1996 to arrange to meet for a taped interview. I would like to schedule the interview at a mutually agreed upon time and place but perhaps during the week that follows .

Accordingly, the interview will focus on the five following research questions. These are:

1. What was the genesis and rationale of the School Achievement Indicators Program?
2. How did the SAIP leaders emerge and what forces facilitated their emergence?
3. What role does the science assessment play in the implementation of the SAIP?
4. How were the framework, criteria, and the instruments of the science assessment designed, constructed, and implemented?
5. What are the implications of the SAIP for education in Canada?

As I previously mentioned, I am a graduate student at the University of Alberta in the Department of Educational Policy Studies. One of the requirements of my program is the completion of a thesis. As a participant, you will be providing me with some data to carry out my study.

Some time after the interview, a transcript of your interview will be provided to you to allow you to review it and make any additions, deletions, and/or changes to ensure that the transcript reflects the intent of your messages. Also, if necessary, could you be available for (a) subsequent interview(s), perhaps by telephone?

Participation in this study is voluntary and you may opt out at any time. As well, I assure you that your participation as well as information shared by yourself will be treated in a confidential manner. With regard to anonymity, sources will not be revealed and pseudonyms will be used.

I sincerely thank you for your time, assistance, and cooperation in this regard.

Yours truly,

Karen Slevinsky
403 427 0010 (work)
403 436 2813 (home)

APPENDICES B1 AND B2

The

Interview

Schedule

and

Follow up

Appendix B1 - The Interview Schedule

Throw-away Questions

1. Questions about the weather, the room, or previous meetings.
2. Questions about the interviewee's day.
3. Comments about setting up the tape recorder, testing it.

Essential Questions

1. How and when was SAIP established?
2. What was Alberta's role in its establishment?
3. How did the development of the SAIP instruments occur?
4. How were the instruments implemented?
5. What was the impact of the SAIP on education in Canada?

Extra Questions

1. What was the role of the other provinces and territories?
2. What role did personnel from the CMEC secretariat play?
3. Likely, what were the belief systems of those involved in the SAIP?

Probing Questions

1. What do you mean by . . .
2. Who was involved?
3. Why did this happen?
4. Do you know more about . . .

Wrap-up Question

1. What are your personal reflections about your specific involvement in the SAIP?
2. What are your personal reflections about your province's involvement in the SAIP?

Appendix B2 - The Follow-up

What was the official name of the committee or team that you worked on, while your involvement with SAIP?

What were your responsibilities as committee or team member?

For what duration did this committee or team exist?

Did you have any involvement with the CMEC or the SAIP prior to this committee or team work?

If the committee or team has been disbanded, do you have any current involvement with the SAIP?

APPENDIX C

Transcript Letter

and

Thank you

June 23, 1996

Dear

Thank you for participating in my study entitled The School Achievement Indicators Program 1988 - 1996, An Historical Survey and Policy Analysis.

Enclosed find a copy of the transcript of our interview that took place on June 10, 1996. Please feel free to review the transcript and make any additions, deletions, and/or changes to ensure that it reflects the intent of your messages. Should you request any changes please contact me by telephone or write your changes on this transcript and mail it back to me before July 15, 1996. **If there are no changes, no reply is necessary.**

Participation in this study is voluntary and you may opt out at any time, however I am expecting that my work is completed by August 1, 1996. As well, I assure you that your participation as well as information shared by yourself has been treated in a confidential manner. With regard to anonymity, sources will not be revealed and pseudonyms will be used, however reference to the titles of positions may be made.

I sincerely thank you for your time, assistance, and cooperation in this regard.

Yours truly,

Karen Slevinsky

403 427 0010 (work)
403 436 2813 (home)

32 Blue Quill Crescent
Edmonton, Alberta
T6J 6C4

Enclosure

APPENDIX D

Acronyms

Acronyms

SAIP	School Achievement Indicators Program, called School Achievement Indicators Project prior to 1990 and called Canadian education indicators or national education indicators prior to 1989
CMEC	Council of Ministers of Education, Canada
ACDME	Advisory Council of Deputy Ministers in Education
OECD	Organisation for Economic Co-operation and Development
CERI	Centre for Educational Research and Innovation
TIMSS	Third International mathematics and Science Study
ESE	Elementary and Secondary Education
NAEP	National Assessment of Educational Progress
CAAT	Canadian Adult Achievement Test
NGOs	Nongovernmental Organizations
PCEIP	Pan-Canadian Education Indicators Program

APPENDICES E1 AND E2

Letters

of

Permission

Devonian Building, West Tower
11160 Jasper Avenue
Edmonton, Alberta
Canada T5K 0L2

November 30, 1994

Ms. Karen Slevinsky
32 Blue Quill Crescent
Edmonton, Alberta
T6J 6C4

Dear Ms. Slevinsky:

Thank you for your letter requesting permission to draw on government materials and personnel to carry out your research study on the School Achievement Indicators Program. I am pleased to grant this permission and request that you provide the department with a copy of your resulting research report.

I expect that the staff in Student Evaluation will be able to offer you considerable support. In addition, if I can be of any assistance during the research, please let me know. I look forward to seeing the results of this interesting study.

Sincerely,

Roger Palmer
Assistant Deputy Minister
Student Programs and Evaluation

cc Frank Horvath
Phill Campbell

252 Bloor West, Suite 5-200, Toronto, Canada M5S 1V5
Telephone 416-964-2551 Fax 416-964-2296 E-Mail CMEC@CMEC.CA

June 6, 1996

Ms. Karen Slevinsky
32 Blue Quill Crescent
Edmonton, Alberta
T6J 6C4

Dear Ms. Slevinsky:

Thank you for your copy of the transcript of our interview that took place in July 1995. Enclosed you will find some minor changes to the transcript.

In regard to your request for permission from CMEC to write your study, of course, the Council supports it. I recognize that my participation as well as the information shared by me has been and will be treated in a confidential manner, and I thank you for respecting that request.

Wishing you every success in your endeavours.

Sincerely yours,

APPENDIX F

1994

Memorandum

of

Understanding

COUNCIL OF MINISTERS OF EDUCATION, CANADA
SCHOOL ACHIEVEMENT INDICATORS PROGRAM

MEMORANDUM OF UNDERSTANDING (MOU)

Introduction

This Memorandum of Understanding (MOU) sets out the overall principles that will guide the School Achievement Indicators Program (SAIP) in all its phases. It will be revised periodically to reflect the evolving nature of the SAIP.

Background

In 1989, the Council of Ministers of Education, Canada (CMEC) initiated the SAIP, the first-ever attempt to arrive at a consensus on the elements of a national assessment. In a Memorandum of Understanding (December 1991), the ministers agreed to measure the achievement of 13-year-olds and 16-year-olds in reading, writing, and mathematics. The choice of the two age groups was made in order to provide for a study of the change in knowledge over time.

In September 1993, at their 64th meeting in Victoria, British Columbia, ministers agreed to the "continuation and expansion of the current School Achievement Indicators Program (SAIP), adding science to mathematics and reading and writing."

Objectives

The SAIP has the following objectives:

- to provide data that will assist each province and territory in making policy decisions, setting education priorities , and planning program improvement, while respecting the autonomy of provinces and territories in matters of education;

- to collect information on the achievement levels in mathematics, reading and writing and science of 13-year-olds and 16-year-olds, that will show how well students of those ages are performing and that will help to determine the effectiveness of the education systems;

- to report in a manner that clearly informs Canadians of the information gathered concerning mathematics, reading and writing, and science, and of the procedures by which the program is conducted, and also provides a possible information base for discussion leading to curriculum improvement at the provincial or territorial level.

In the SAIP, 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 is essentially a measure of how well each jurisdiction's education system is doing and does not replace individual student assessment, which is the responsibility of teachers, schools, boards, and ministries of education. Similarly, no attempt is made to compare schools or school districts.

Management

In order for a national assessment program to serve the purposes of education, it is imperative that the provinces and territories direct the development of the SAIP and that all jurisdictions have an equal voice in its direction. These principles are reflected in the funding and administrative arrangements for the SAIP. Moreover, all key decisions related to the SAIP are subject to review by the ministers.

The CMEC Secretariat is responsible for the overall management of the Program, coordinating activities related to the planning and financing of the SAIP, and providing the focus, at the national level, for the promotion of the Program. Other responsibilities include providing advice on the administration of the Program, and assisting the various consortia and CMEC committees in carrying out their tasks.

It is the intention of the Council that the administration and marking component of the assessment be contracted out. A consortium of provinces will be established to develop assessment materials for the Science component. Another consortium of provinces will be established which will be responsible for revising the assessment instruments for mathematics and language, in preparation for subsequent assessments. The role of the consortia, participating jurisdictions and the CMEC Secretariat remain to be determined once the operational details of the administration model have been approved by the Council, the goal being that such approval be given by March 1, 1995.

The management structure for the SAIP is as follows:

- Policy Advisory Committee (PAC): Composed of the deputy ministers responsible for elementary and secondary education, its responsibilities include making decisions regarding policy and administration matters, and developing appropriate recommendations for the consideration of the CMEC.

- Administration Management Team (AMT): As agreed to by the CMEC in September 1992, the AMT is comprised of the Chair of the ACDME, the Director General of the CMEC, and the Chair of the Report Development Group (RDG) and three members of the RDG. Its responsibilities include acting as the first-level policy reference point on administration, marking and reporting issues; developing recommendations for the Policy Advisory Committee; selecting and overseeing the work of the report writer.

- Report Development Group (RDG): Composed of representatives from all jurisdictions, its responsibilities include developing framework for the final reports (public and technical); working with the writer to ensure that these reports are in keeping with the agreed-to framework; making recommendations to the Policy Advisory Committee.

- Provincial and territorial coordinators: One contact person from each jurisdiction, chosen by the deputy minister, forms this group which is responsible for coordinating the administration of the SAIP in his/her jurisdiction.

- Science Development Team: Composed of representatives from the four consortium partners, namely Alberta, Saskatchewan, Ontario and New Brunswick (francophone), the Team is responsible for developing the science assessment instruments.

Financing

Based on program expenditure projections developed by the CMEC Secretariat in consultation with provincial and territorial officials, annual and long-term budget forecasts for the SAIP will be presented to the CMEC. The SAIP budget will be prepared separate from that of the CMEC and monetary support will be sought from other than provincial and territorial ministries responsible for education. In soliciting support for outside sources, it is noted that:

- the outcome of any agreement will need to respect the objectives and goals for the development of education indicators, as identified by the CMEC, ensuring that provinces and territories direct the development of the SAIP and that all jurisdictions have an equal voice in its direction;

- the involvement of funding partners could include participation at meeting of the Science Development Team, provincial and territorial contacts, the Report Development Group and any other technical-level meetings and public advisories. As well, it could include access to the reports and results of the development, administration and reporting stages. However, because of the secure nature of the assessment items, access to these would need to be restricted.

Consistent with the decision taken by the CMEC at its 65th meeting in February 1994, the following formula will be used to calculate provincial and territorial direct contributions to the long-term funding of the SAIP: the amount that each participating jurisdiction will contribute to the SAIP will be calculated by first determining the difference between the per capita plus per participating student estimates and the per capita estimates; then, half of the difference will be applied as a correction to the per capita estimates. In addition to any direct cost implications for the SAIP through the CMEC by provincial and territorial authorities, in-kind contributions and services will be committed by participating jurisdictions to support the SAIP in individual jurisdictions.

Administration of Assessments

The achievement information in the different subjects requires the testing of groups of students. The plan is to administer the tests in April of each year according to the following schedule:

Mathematics	Reading and Writing	Science
1993	1994	1996
1997	1998	1999

No full testing is planned for 1995, to allow for field testing of the questions for science.

Guiding Principles

In order to facilitate the development and reporting of the achievement indicators, the members of the CMEC agree to the following points.

1. To the best of their abilities, all provinces and territories will actively participate, in whole or in part, in all phases of the program.

2. A commitment is made to ensuring that the assessment techniques respond to current pedagogical practice. The assessment instruments will include some examination of students' best work in writing, students' personal/reflective responses to reading, student performance in mathematics problem-solving, and student achievement on a practical science assessment.

3. The development of the assessment instruments will incorporate a process that ensures that they are as much as possible in keeping with the curriculum requirements and orientations of the participating provinces and territories. The following procedures will be used to ensure that the assessment instruments are free from cultural and gender bias and stereotyping and that all sectors of the population are treated fairly:

   - specialists in recognizing stereotypes and bias will review all assessment materials intended for student use;
   - each jurisdiction will establish procedures for reviewing potential test items for relevance to its curriculum, for gender and culture bias or stereotyping, and for other necessary quality and validation components;
   - student and teacher feedback will be collected during the field testing of new items;
   - appropriate means will be used by each jurisdiction to ensure parental and interest group feedback;
   - statistical analyses will be conducted on the results of the field test, in order to detect bias;
   - any items in which the wording or content show bias or stereotypes will be discarded.

4. A sample of 13- and 16-year-old students will be selected in each province and territory, through a process that is sensitive to, among other variables, the demographic make-up of the jurisdiction. A jurisdiction may decide to draw core samples for both official languages.

5. For each assessment, a framework will be developed that will allow for the inclusion of additional sample groups to accommodate the different interests of each province and territory. Additional resources needed to meet the special requirements of individual jurisdictions will be the responsibility of those jurisdictions.

6. The range of responsibilities of participating jurisdictions regarding administration, as agreed to by the CMEC at its September, 1992 meeting, are as follows:

   - Designating appropriate ministry personnel to coordinate administration in the province/territory, review and analyze instruments and reports, analyze and interpret results for the province or territory.
   - Selecting schools and students for the sample.
   - Ensuring the provision of any training and information that coordinators, principals, teachers required to administer the instruments in the schools.

- Distributing the copies of the test instruments in the province/territory prior to administration, collecting the instruments upon completion, and returning to marking centres.
- Printing costs - testing materials; provincial and territorial handbook; school coordinator's handbook; promotion materials for teachers, parents, etc.; copies of final report (public and technical) required by each province for distribution to schools, ministry officials, NGOs.
- Distributing information materials on the SAIP to administrators, teachers, trustees, parents, NGOs, and the press in the province/territory.
- Ensuring the quality of test administration.
- Consulting with provincial NGOs and others before the testing and after the results are received and distributing information and reports on the results.

7. The process for report development will include a mechanism for provincial and t territorial approval. In addition to the technical report, if it so desires, a province or territory can ask to receive its results. However, all requests concerning the results for another participating province or territory must be made directly to that jurisdiction. The use and dissemination of results (both from the public report and the technical report) are set out in a protocol or statement that users outside the ministry of education will be asked to sign.

Development of the Science Component

The Consortium Agreement, agreed to by the Science Consortium partners, sets out the parameters for the development of the Science component of the SAIP. It clearly outlines the responsibilities (timelines, deliverables, consultative process, etc.) of each Consortium member and of the CMEC Secretariat in this process.

June 17, 1996

Council of Ministers of Education, Canada
252 Bloor West, Suite 5-200,
Toronto, Canada
M5S 1V5

Dear

I am a graduate student at the University of Alberta in the Department of Educational Policy Studies. I was a member of the SAIP Science Development Team. I am nearing the completion of a thesis entitled The School Achievement Indicators Program 1988 - 1996, An Historical Survey and Policy Analysis.

To enhance my thesis and increase its usefulness, I would like to include a copy of the Memorandum of Understanding for the School Achievement Indicators Program, dated May 20, 1994. I am requesting permission to reprint this MOU and add it to my thesis as an appendix. To be exactly clear about the document to which I am referring, I am enclosing a copy.

Thank you for your attention to this matter and also for the assistance provided to me by the CMEC Secretariat in other related matters. I look forward to your reply.

Yours truly,

Karen Slevinsky
32 Blue Quill Crescent
Edmonton, Alberta
T6J 6C4

cc Ms. Dianne Pennock
National Coordinator SAIP

ENCLOSURE

APPENDICES G1, G2, G3,

AND G4

The

SAIP Science Assessment

Instrument

Blueprint

Gr							
B			A02				
CHE			A08				
EART			A12				
PHYSI	CP	A03	A05				
NATURE	NS	A10	A11				
SCIENCE	ST	A04	A06				
Groupe 2							
CONCEPTUAL	C	A01	A02	A03	A06	A11	A12
PROCEDURAL	P	A04	A07	A08	A10		
USE	U	A05	A09				
Groupe 3							
LEVEL3	3	A01	A02	A03	A04	A05	A06
		A07	A08	A09	A10	A11	A12

WIN1.1SAIP / SCIENCE -- BOOKLET B 1996-07-31 08:14

[4] Regroupements par thèmes

Groupe 1

BIOLOGY	CB	B27	B37	B38	B40	B41	B44
		B64	B01	B14	B26		
CHEMISTRY	CC	B06	B15	B19	B20	B34	B42
		B51	B54	B55	B59		
EARTH	CE	B02	B10	B28	B36	B45	B47
		B48	B60	B65	B66		
SCIENCE	CP	B18	B21	B22	B23	B25	B49
		356	B57	B58	B61		
SCIENCE	NS	B04	B05	B09	B11	B12	B29
		B30	B31	B32	B33	B39	B43
		B63					
SCIENCE	ST	B03	B07	B08	B13	B16	B17
		B24	B35	B50	B52	B53	B62
		B46					

Groupe 2

CONCEPTUAL	C	B08	B10	B12	B17	B18	B19
		B20	B22	B23	B24	B27	B29
		B35	B36	B37	B40	B41	B42
		B43	B45	B47	B49	B50	B51
		B52	B53	B56	B57	B59	B62
		B64	B65	B14	B46		
PROCEDURAL	P	B03	B04	B06	B11	B13	B15
		B16	B28	B30	B31	B32	B39
		B55	B61				
USE	U	B02	B05	B07	B09	B21	B25
		B33	B34	B38	B44	B48	B54
		B58	B60	B63	B66	B01	B26

Groupe 3

LEVEL1	1	B04	B05	B06	B07	B13	B17
		B19	B20	B21	B22	B28	B29
		B36	B38	B39	B40	B43	B44
		B49	B53	B55	B56	B65	B66
		B01	B46				
LEVEL2	2	B03	B08	B09	B12	B15	B16
		B23	B24	B25	B27	B31	B32
		B33	B37	B41	B42	B45	B47
		B48	B50	B51	B54	B57	B60
		B61	B26				
LEVEL3	3	B02	B10	B11	B18	B30	B34
		B35	B52	B58	B59	B62	B63
		B64	B14				

WIN1.1SAIP SCIENCE / BOOKLET C 1996-07-29 09:06

[4] Regroupements par thèmes

Groupe 1

BIOLOGY	CB	C07	C16	C35	C36	C37	C38
		C54	C56	C57	C47		
CHEMISTRY	CC	C08	C11	C12	C13	C17	C34
		C50	C51	C52	C53		
EARTH	CE	C01	C04	C06	C18	C22	C24
		C25	C55	C63	C41		
PHYSICS	CP	C09	C10	C15	C20	C28	C29
		C49	C60	C61	C62		
NATURE	NS	C02	C05	C21	C30	C31	C32
		C39	C44	C46	C64	C65	C66
		C33					
SCIENCE	ST	C03	C14	C19	C23	C26	C42
		C43	C45	C58	C59	C27	C40
		C48					

Groupe 2

CONCEPTUAL	C	C01	C03	C04	C08	C10	C12
		C14	C16	C17	C18	C19	C20
		C23	C24	C25	C34	C35	C36
		C42	C43	C45	C49	C50	C53
		C55	C57	C58	C59	C62	C64
		C66	C27	C33	C47		
PROCEDURAL	P	C02	C05	C09	C11	C30	C31
		C32	C37	C39	C44	C51	C54
		C60	C61				
USE	U	C06	C07	C13	C15	C21	C22
		C26	C28	C29	C38	C46	C52
		C56	C63	C65	C40	C41	C48

Groupe 3

LEVEL_3	3	C01	C02	C08	C13	C14	C15
		C21	C22	C26	C36	C44	C45
		C49	C47				
LEVEL_4	4	C03	C04	C06	C09	C11	C16
		C17	C18	C20	C29	C30	C32
		C34	C37	C38	C39	C42	C46
		C51	C55	C56	C59	C61	C27
		C33	C48				
LEVEL_5	5	C05	C07	C10	C12	C19	C23
		C24	C25	C28	C31	C35	C43
		C50	C52	C53	C54	C57	C58
		C60	C62	C63	C64	C65	C66
		C40	C41				

Science / Practical -- Training **1996-08-02 20:32**

[4] Regroupements par thè:

Groupe 1

LEVEL_1	1	T1_5 T7_2	T2_3	T3_1	T4_1	T5_1	T6_5
LEVEL_2	2	T1_7 T7_5	T2_5	T3_3	T4_3	T5_2	T6_6
LEVEL_3	3	T1_8 T7_6	T2_6	T3_4	T4_5	T5_4	T6_7
LEVEL_4	4	T1_9 T7_7	T2_7	T3_5	T4_6	T5_5	T6_8
LEVEL_5	5	T1_10 T7_8	T2_8	T3_6	T4_7	T5_6	T6_9

Groupe 2

OBSERVATION	A	T1_5 T7_2	T1_7 T7_5	T2_3	T2_5	T6_5	T6_6
CONCLUSION	B	T1_8	T4_5	T5_4	T6_7		
EVALUATION	C	T1_9	T3_5	T4_6	T5_5		
LIMITATION	D	T1_10	T3_6				
VARIABLES	E	T2_6	T2_8	T3_3			
ERRORS	F	T2_7	T6_8				
PURPOSE	G	T3_1	T4_1	T5_1			
PATTERNS	H	T3_4					
PROCEDURE	I	T4_3	T5_2	T7_8			
RELIABILITY	J	T4_7	T5_6				
INTERPRETE	K	T6_9					
GRAPH	L	T7_6					
MATH	M	T7_7					

Curriculum Vitae

Personal Data:

Karen Slevinsky
32 Blue Quill Crescent
Edmonton, Alberta
T6J 6C4

Education:

Enrolled in the Master of Educational Administration program at the University of Alberta, fall 1996 expected completion

Enrolled in the B.Sc. degree program, University of Waterloo, January 1997. expected completion

1981, Diploma of the Faculty of Education, Curriculum and Instruction, University of Alberta

1978, Bachelor of Education, University of Alberta

Awards:

Nominated for the 1994-95 Edmonton Public Schools Staff Merit Award

1994 Prime Minister's Award for Teaching Excellence in Science, Technology and Mathematics, National Recipient

Professional, Teaching & Administrative Work:

1995 - present, Examination Manager, Biology 30, Student Evaluation Branch, seconded to Alberta Education from Edmonton Public Schools

1994 - 1995, Science Specialist, School Achievement Indicators Program, Council of Ministers of Education, Canada, seconded to Alberta Education from Edmonton Public Schools

1992 - 1994, Science Department Head and science teacher, Jasper Place High School, Edmonton Public Schools

1981 - 1992, biology, chemistry and International Baccalaureate biology teacher, Harry Ainlay High School, Edmonton Public Schools

1980 - 1981, grade seven science teacher, Vital Grandin School, St. Albert Catholic Schools

1978 - 1980, junior high social studies and language arts teacher, Mistassiniy School, Northland School Division

Extracurricular Activities:

Edmonton Public Schools

District: Event Manager, Science 10 "Passing Through" event for the Senior High Science Olympics, March 12, 1994

Attended the December 1, 1993 Study Session: Developing, Enhancing, and Evaluating Staff Performance

Participated in the Edmonton Experience V, November 19-21, 1992, Centre for Education, Edmonton Public Schools

Twinned with the Academic Challenge teacher and class at Greenfields Elementary School; presented labs and hosted an exchange of students, 1991-92

Participated in the Edmonton Public Schools Leadership Seminars, December 1990 - March 1991

Event Manager for the Senior High Science Olympics for 1984, 1985, and 1987

School: **School Wide Projects**

January and June Examination Week Policy, spring 1994

A WHMIS safe school, fall 1993

Science Olympics: coached and assisted with the coaching of two or more teams of students so as to encourage student participation in science activities, each year since the inception of Senior High Science Olympics, 1984

Writing Projects

Discovering the Abiotic Factors Associated with Snow Accumulation, February 1994

JP 500, Mouse Trap Car Project for Science 10, January 1993

Fetal Pig Dissection Laboratory and Photographs

Plant Ecology Laboratory

Drosophila Giant Chromosome Extraction and Photomicrographs

A Student's Guide to Identifying Common Trees, Shrubs and Herbs in Whitemud River Valley

Field Trips numerous

Alberta Education Curriculum Standards Branch

Teacher representative on the Secondary Science Advisory Committee for Curriculum Standards, October 1993 - October 1995

Presented at the Alberta Education 30-level Information Sessions, March 18, 19, 24, 25, 26, 1994 in Grande Prairie, Airdrie and Lethbridge; "Evaluating Skills in the Classroom: Problem Solving Model in Action"

Wrote a portion of the "Exemplary Performance Assessment Laboratories" document as directed by Raja Panwar, November 1993 - March 1994

Coordinated the writing of the four unit examinations and the final examination for Science 20, March - May 1992

Wrote questions for the Unit 1 examination for Science 10; coordinated the compilation of questions into a Science 10 Final Examination, January - March 1992

Organized and updated the Assessment and Evaluation section of the Teacher Resource Manual, January - March 1992

Presented 10 laboratories for Science 10, Unit 1; a workshop for teachers at the U of A, August 1991

Wrote Sample Assignments for Science 10, Unit 1 Teacher Resource Manual, May - August, 1991