The British Columbia Assessment Program: An Overview

Jerry J. Mussio and R. Nancy Greer
british columbia ministry of education

This paper examines the organization and impact of the British Columbia Assessment Program. A series of historical events which have influenced the development of the B.C. Assessment Program are discussed. The paper reviews specific follow-up actions to the assessment program at both the provincial and school district level. Difficulties encountered in the design and implementation of assessment programs are examined.

Cet article se penche sur l'organisation et les répercussions du programme d'évaluation de la Colombie-Britannique. Il présente une série d'événements historiques qui ont eu un effet sur son évolution. L'article expose également des mesures précises auxquelles le programme a donné lieu tant au niveau provincial qu'au niveau local du système scolaire. Il examine enfin les difficultés qui se posent dans la conception et la mise en application des programmes d'évaluation.

Among the important topics debated throughout the country during the 1970s is the question of accountability in education and, more specifically, government involvement in the area of student testing. Discussion of this issue undoubtedly has been influenced by a flurry of testing activity in the United States and extensive publicity over the state of literacy throughout North America.

Reports of poor achievement have added substance to the view that educational practices such as the use of provincial examinations, which succeeded and appeared appropriate for the past, can satisfy new demands for accountability and better management of provincial educational systems. Looking backward for solutions of this type, however, can succeed only when social conditions and educational aspirations of the general public remain relatively stable.

In the first part of this paper we explore how some of these conditions and aspirations have changed and how, in our judgment, they have influenced the development of the British Columbia Assessment[1] Program. The second part of the paper reviews the organization and impact of the B.C. program. In the last part of the paper, we examine a series of issues which need to be considered in the design and implementation of provincial assessment programs in general.

EXAMS AND RETENTION RATES

In 1973, the last set of regular government examinations[2] was administered in B.C., a trend characteristic of most provinces at that time. This

move away from provincial examinations was a major event in that it represented the end of a 100-year-old tradition. In order to appreciate the context of the B.C. Assessment Program, it is useful to examine a number of events which preceded the shift which took place in 1973.

Traditionally, provincial examinations in Canada have been used to determine who would be admitted to the next grade or who would be admitted to university. Egerton Ryerson, a pioneer of educational life in Ontario, was primarily responsible for the introduction of the examination system into his native province well over 100 years ago. From Ontario its influence spread throughout Canada to British Columbia.

In British Columbia, government examinations were first used to admit students to the system of high schools which had been established in 1874. Students failing these examinations would repeat Grade 8 or leave school for the job market — not a bad alternative at that time. In high school, each year ended with a set of examinations which again were used to determine promotion to the next grade and, finally, to determine admission to university.

The early part of the twentieth century saw teachers and parents urge a relaxation of the rigid examination policies in the province. The Putnam-Weir Report (1925), which was to affect the course of education in B.C. over the next 50 years, recommended that as the standards of teaching and qualifications of teachers improved, students should be permitted to enter higher levels of schooling solely on the recommendation of their teachers.

By 1931, teachers were given the authority to determine promotion through to Grade 11. By 1937 students with a C+ average from accredited schools were being admitted to university without sitting for a departmental examination. The Putnam-Weir recommendation dealing with promotion was ultimately carried out in 1973 when the last set of government examinations was administered in British Columbia.

The effects of changes in government examination policies coupled with increased emphasis in society on the importance of higher education have led to dramatic changes in student retention rates over the past 50 years. In 1921, for example, 7,749 students were enrolled in Grade 8 in B.C. Of this group 3,996, or 52% of the Grade 8 population, were admitted to Grade 9 (Table 1). Three years later, in 1924 (Table 2), a total of 1,960 of the original group of students were enrolled in Grade 11, a retention rate of 25% between Grade 8 and Grade 11. (Grade 11 at that time was the last year of a three-year high school program, and Grade 12 was the equivalent of a first year university program.) In 1937, the retention rate between Grade 8 and 11 had risen to 40%. By 1952 the rate climbed to 62%. And by 1972, 92% of those students who were enrolled in Grade 8 in 1969 were enrolled in Grade 11 in British Columbia schools.

TABLE 1

Admission to Grade 9, British Columbia, 1920-23[a]

Year	*Enrolled Grade 8*	*Admitted to Grade 9*[b]	*% Who Passed Grade 8*
1920	6,780	3,362	49.6
1921	7,749	3,996	51.6
1922	8,608	3,585	41.7
1923	8,897	4,579	51.5

[a] Putman-Weir Report (1925), p. 264.

[b] This includes those who passed on recommendation.

TABLE 2

Retention Rates. Grades 8-11, 1921-1969[a, b]

Year	*Grade 8*	*Year*	*Grade 11*	*% of Grade 8 Enrolment*
1921	7,749	1924	1,960	25
1934	11,323	1937	4,476	40
1949	11,931	1952	7,373	62
1969	39,265	1972	36,265	92

[a] B.C. Education Annual Reports (1921-1973).

[b] For example, of the 7,748 students who started in Grade 8 in 1921, 1,960 or 25% were registered in Grade 11 in 1924. Note that these data were not adjusted for movement into and out of the system.

A recent Ontario study (Traub et al., 1976) points out that one half of today's high school students would not have been in high school 15 years ago. To dramatize this point, if half of the high school population were to leave school tomorrow, B.C.'s unemployment rate would jump from 8.5% to about 14%.

Not only has there been a change in the numbers of students attending higher levels of schooling, but there has been a change over time in how literacy is defined. And changing definitions result in changing estimates of a population's literacy. If writing one's name were what was meant by literacy, for example, society would not be worried that illiteracy was a national problem. Yet until the nineteenth century this skill was not shared by the majority of the population in North America and even in the more developed nations of Europe (Schofield, 1968).

Resnick (1977) suggests that perhaps society perceives itself as having an unacceptable literacy level because it is applying a criterion that

requires, at a minimum, the reading of new material and the gleaning of new information from the material. He claims that this high literacy standard is a relatively recent one as applied to the population at large, and much of the present difficulty in meeting the literacy standard society is setting for itself can be attributed to the relatively rapid extension to large populations of educational criteria that were once applied to only a limited elite. Hence, one can observe the problems high schools are now experiencing in dealing with students who at one time likely would have dropped out of school and who now are confronting problems with what are now considered to be basic skills.

In short, where higher education was once thought appropriate for a select few, it appears that throughout Canada we have moved from a period characterized by scarcity of educational opportunity to an era of abundance and choice, where all parents now expect their children to graduate from high school, at the very least.

RESPONSIBILITIES OF GOVERNMENT

The shift from a selective form of education to a system where all students are expected to complete what were once thought to be advanced levels of education raises a number of fundamental issues for a provincial government. These can be summarized as follows:

1. If society has accepted the principle of universal education, then it would appear that a provincial government, by virtue of the mandate outlined in the BNA Act, has a fundamental responsibility to indicate in some detail the basic learnings which the vast majority of students should be expected to acquire. Phrased differently, a provincial government should be in a position to describe the fundamental goals and objectives of the public school system.

2. It would then seem essential for government to determine to what extent the principle of universal education is being realized. In other words, government should be able to respond to the question: To what extent are all students in the province acquiring the basic outcomes of the provincial school system?

3. Once this latter question has been addressed, government must assume responsibility for ensuring that a greater proportion of our youth are acquiring these basic learnings. A provincial government, therefore, should also be able to respond to the question: What is being done to ensure that even more of our youth achieve an acceptable level of education? Response to this question could mean changes to the existing curriculum, better selection of learning materials, improved professional development programs for teachers, a reallocation of resources, and so on.

Many will argue that, in today's world, a central authority such as a ministry of education has no business being involved in assessment or even curriculum. However, given the nature of the BNA Act, wherein provincial governments have been granted the authority and responsibility over education, and given the fact that provincial governments now are accountable for billion dollar educational budgets, it is difficult to see how this involvement can be avoided. In our judgment, serious consideration of the three points raised above should be viewed as an essential component of responsible educational management at the provincial level.

An important issue that is often debated is whether or not a provincial government should be involved in measuring and reporting individual student achievement. If a province wishes to serve the purposes of both individual student assessment and provincial level program monitoring, every pupil must be tested with the same test. Whether this approach will produce optimal data for both purposes will depend on the variability of instructional objectives and in pupil performance levels. The greater the span of objectives and in student performance, the longer the test must be to capture the full range of the curriculum and achievement. Shepard (1978) argues that at some point it becomes more reasonable to administer efficiently a comprehensive matrix-sampling test for provincial purposes and leave districts or teachers free to administer tests for individual pupil diagnosis. We agree with this position.

THE B.C. ASSESSMENT PROGRAM

Following the termination of provincial exams in 1973, the British Columbia Ministry of Education, Science and Technology formed a Joint Committee of Evaluation to advise on the development of a long-range assessment plan for the province. The committee, consisting of teachers, school trustees, superintendents, university personnel, and ministry staff, met for a period of about two years. The rationale developed by this committee has formed the basis of the B.C. Assessment Program (Conry et al., 1975). The fundamental purpose of the assessment is to provide information as to what and how students are learning on a province-wide basis. Each assessment includes five basic components, as outlined in Figure 1.

Goals Survey

Information is gathered from teachers and members of the public regarding the suitability of various goals and learning outcomes. This information can serve two purposes: It can help to identify priority skill areas which should be assessed and it can also provide curriculum revision committees with feedback on the appropriateness of existing curricular objectives or the desirability of adding other important outcomes.

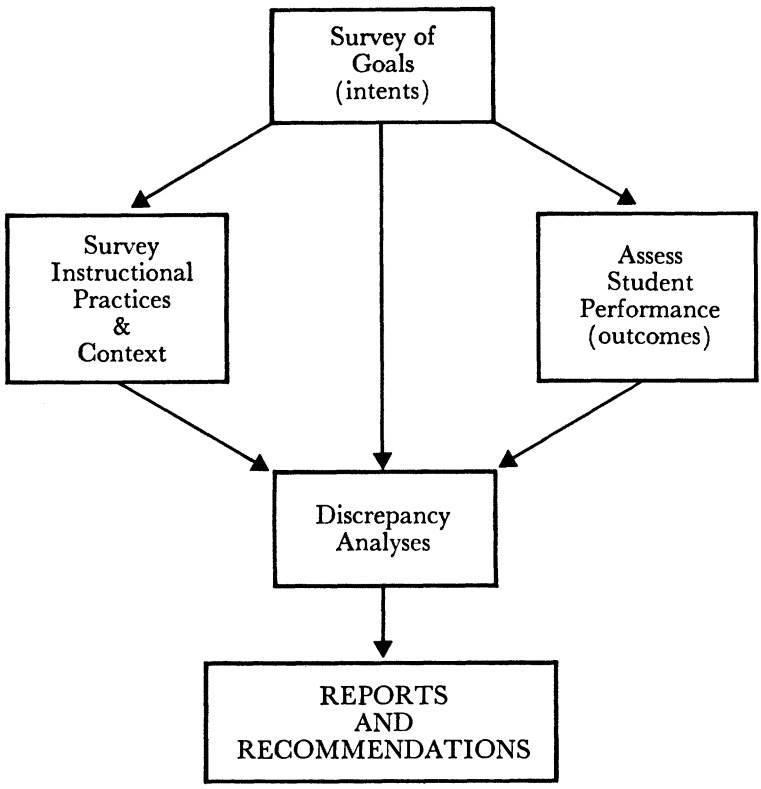

FIGURE 1

B.C. Assessment Model

Student Performance

The B.C. Assessment Program attempts to address the question: To what extent are all students achieving the basic objectives of the public school system? An important question that appeared to be addressed in the former provincial examination system was: What percentage of students should be admitted to higher levels of schooling? The distinction between these two questions has been critical in the design of the provincial assessment program in British Columbia.

Since the results of the B.C. assessment are reported for groups and not for individual students, it is not necessary that every pupil at each grade level take the same set of exercises. Student achievement, therefore, is generally assessed using matrix sampling procedures with group results being reported at the provincial and/or district level. Individual students

and schools are not identified and the decision as to whether district results are to be provided is dependent on the costs of a particular assessment. A writing assessment, complete with open-ended exercises, for example, is much more expensive than a mathematics assessment consisting of multiple-choice questions. In the case of a writing assessment, then, only provincial results are reported, whereas mathematics results are reported at the district level as well.

Instructional Practices

A survey of instructional practices in the schools is usually conducted as part of each assessment. This includes a survey of program organization in the schools, texts and instructional materials, teacher background and training, methods of evaluation, and classroom activities. The inclusion of this type of survey in the assessment program is based in part on the notion that this information will help to explain discrepancies between expected and actual levels of student learning.

It is generally recognized that certain learning outcomes cannot be assessed given the current state of the art. We can, however, assess classroom practices and availability of resources that educators consider necessary to contribute to those experiences. For example, we assume that exposure to good music and good literature contribute to the intellectual and aesthetic development of children. At present it is very difficult, if not impossible, to assess the students' aesthetic development, but we can attempt to assess whether or not children in the schools ever hear good music or read good literature. It is to this end that a survey of instructional practices can also contribute.

Discrepancy Analysis

The assessment test results are reported in terms of the percentage of students successfully answering each exercise. Panels of teachers and members of the public are then called upon to judge the results of each exercise. Judgments of strength and weakness are based not only on the percentage of students who have correctly responded but also on the perceived importance of the concept tested by the item. Whenever district results are reported, district staff are asked to involve teachers and members of the public in the interpretation of the district data. Extensive guidelines for the interpretation of district results are provided for each assessment (Ministry of Education, 1976; 1977; 1978).

Reports and Recommendations for Change

The final phase is the production of assessment reports which focus on the generation of recommendations for follow-up action. These recommendations, which are derived from all of the data collected in the

assessment, are directed at specific groups including the ministry, teacher training institutions, school districts, schools, and teachers.

Teams contracted by and external to the ministry, which consist of teachers and university specialists, are responsible for the development of all assessment instruments and final reports. Each contract team works with an advisory committee which serves as a sounding board for the project. The assessment staff of the ministry is responsible for the overall management of each assessment. A technical agency works with the contract team and the ministry to handle all administrative and data analysis aspects of the assessment.

It is intended that these assessment reports, and the B.C. Learning Assessment Program in general, will be used to:

1. monitor student achievement over time;

2. assist curriculum developers at the provincial and local levels in the process of improving curriculum and developing suitable resource materials;

3. provide information which can be used in determining the allocation of resources at provincial and district levels;

4. provide direction for change in teacher education and professional development;

5. provide direction for educational research;

6. inform the public of some of the strengths and weaknesses of the public school system.

Assessments completed to date and a tentative schedule of future assessments are outlined in Table 3.

THE IMPACT OF A PROVINCIAL ASSESSMENT

The value of any provincial assessment activity must ultimately be judged on the basis of its impact on the day to day conduct of the province's educational programs. In other words, to what extent are the purposes of the program being met?

Provincial Follow-Up

In the four years of assessments completed to date in British Columbia, there have been, in our view, a number of examples of positive follow-up at both the provincial and the school district level. For example, the results of the first assessment (Evanechko et al., 1976) served as a base for the revision of curriculum for the elementary and secondary language arts programs. These new programs were implemented starting in fall of 1979.

TABLE 3

B.C. Assessment Timetable

Year	*Content Area/ Grade Level*	*Smallest Reporting Unit*
Pilot 1976	Reading 4	District
	Writing 8, 12	Province
1977	Reading 8, 12	District
	Math 4, 8, 12	District
	Social Studies 4, 8, 12	Province
1978	Writing 4, 8, 12	Province
	Science 4, 8, 12	District
1979	Physical Education 3, 7, 11	Province
1980	Reading	District
	Kindergarten	Province
1981	Mathematics	District
1982	Social Studies 4, 8, 12	District
	Science 4, 8, 12	District
	Special Programs	Province
1983	Industrial Education	Province
1984	Reading 4, 8, 12	District
	Writing 4, 8, 12	Province

Other examples of follow-up in the area of curriculum have included modifications in the provincial mathematics guide to include a stronger consumer mathematics component, an area of weakness identified province-wide in the 1977 mathematics assessment (Robitaille et al., 1977). The extensive survey of goals that formed a major part of the 1977 social studies assessment (Aoki et al., 1977) has served as a substantial information base from which to begin a major revision of the secondary curriculum in this area, scheduled for completion in the spring of 1981. Similarly, results of a recent science assessment (Hobbs et al., 1978) which pointed to concern over aspects of the science program have led to a number of follow-up actions, including, for the first time, the formation of a K-12 provincial science advisory committee.

Other forms of follow-up not strictly related to curriculum have also been undertaken by the ministry. Discussions with district representatives indicate a strong demand for provincially prepared tests and test items

keyed to the specific course objectives and accompanied by provincial norms. Twelve such tests in mathematics have been developed to date and are available to teachers on request. As well, kits of provincially normed writing exercises and marking scales have been made available to all schools recently in response to recommendations made in the report of the provincial assessment of written expression (Conry & Rodgers, 1978). Similar developments are currently underway in both senior secondary science and physical education.

It will be recalled that as well as providing input on matters of curriculum, major purposes for the B.C. Learning Assessment Program include the provision of information to guide the ministry in decisions related to resource allocation and in the identification of areas within the province which may be experiencing less than satisfactory levels of performance. In the areas of reading, mathematics, and science separate reports of performance were generated for each of the 75 school districts in the province.

Following the first assessment in reading which was reported in 1976, the Grade 4 reading performance of five districts fell substantially below the level established as acceptable by provincial interpretation panels. This judgment of weakness was confirmed by school officials in these districts and resulted in co-operative efforts in the form of personnel and resource assistance by the ministry to address the identified problem areas. In one district this follow-up took the form of a program of professional development for teachers in the teaching of reading, the purchase of instructional support materials, and the implementation of a testing program to identify needs of individual students. In another of the districts, further inquiry led to the recognition that the problem was of a socio-economic nature that could not be addressed by educational means alone. An inter-ministerial committee of Education, Human Resources, and Attorney-General personnel was established to deal with the non-school social issues which appeared to be related to the less than satisfactory performance of the school children within the district.

Regression analyses conducted by the ministry (Greer, 1978) to identify those factors most predictive of performance across the districts led to a major initiative on the part of the Native Indian Education Branch within the ministry. The results of this initiative over the next several years, it is hoped, will be a substantial increase in curriculum materials specifically for teachers of these children.

District Level Follow-Up

Several cases of positive follow-up in response to assessment reports have also occurred at the district level. As mentioned earlier, a number of the assessments completed to date have provided individual reports of group achievement to each district, along with extensive guidelines to

complete a local interpretation process. Reports from local interpretation panels, and the recommendations generated, are sent to the ministry and are discussed during annual regional meetings held with senior district personnel and Assessment Branch staff members.

The initial skepticism of districts that greeted the first assessment has dropped substantially in the last two years, due largely to considerable time spent by assessment staff meeting with administrators and teachers throughout the province. As a consequence, the vast majority of districts now complete extensive interpretations of their own results, develop reports for their school boards, and implement some form of follow-up. The following examples are taken from district reports submitted to the ministry.

In Nanaimo, a district of just over 10,000 pupils, recommendations tabled with the school board following the 1977 reading assessment resulted in the appointment of an English co-ordinator, in the establishment of a standing district language arts committee, and the launching of a series of inservice workshops for all teachers (Nanaimo, 1977). In Victoria the same assessment led to the recommendation to the school board and subsequent approval for the establishment of reading centres in all 15 secondary schools (Victoria, 1979). Funds have since been approved by the board to support the second year of operation of these centres which are geared to provide reading instruction to high school students, either individually or as classroom groups.

In a large urban district with a sizeable and steady influx of children for whom English is a second language, the reading assessment results led to the institution of an annual program in reading for all students enrolled in Grades 10 and 11 to identify students requiring tutorial assistance. Still another district's response to less than fully satisfactory reading results was to arrange a co-operative venture with one of the local universities to offer an off-campus course in developmental reading after school for its junior secondary teachers. In an isolated district with limited community resources and a large native population, a recreational reading period for 30 minutes each day was initiated for all students in its secondary school. The district reported noticeable changes in attitudes of the students as evidenced by the marked increase in use of library facilities in the six months since this plan was instituted.

Two comments frequently made by districts should be noted here. One is that the assessments have not so much allowed for entirely new initiatives to be launched, but rather have served as the necessary catalysts to bring about school board approval for projects that had been pending. The second frequent observation is that the local interpretation process has forced teachers, trustees, and parents to communicate, often for the first time, about actual data describing student achievement. As well, the assessments have often led to increased communication among professionals within the district across subject disciplines and grade levels.

AREAS OF CONCERN

From the preceding examples of follow-up it can be seen that assessment in British Columbia has had some impact on the conduct of education. It would be misleading to suggest, however, that the program is, to our way of thinking, unequivocally positive and without flaws. Indeed, the experience over the past several years has served to bring into focus a number of issues which we consider to be problematic not only to British Columbia but to the conduct of assessment programs in general.

In discussions with other jurisdictions we find these same problem areas to be shared by others. A paper of this sort would not be complete without mention of some of the more fundamental and recurrent of these problem areas. It is our intent in this last section to highlight these issues, not because we have found satisfactory resolution, but rather because they remain open questions in need of resolution if large-scale assessments are to fulfill their designated purposes in the provincial or local setting.

These issues can be classified into four general categories: measurement and instrumentation; interpretation; communication; utilization.

Measurement and Instrumentation

It is a well-accepted first step in educational measurement to establish a clear and precise definition of what is to be measured. The work of individuals such as Hambleton & Eignor (1979), Popham (1978), and Millman (1978) has produced procedures and "rules of thumb" for domain specification and for ensuring adequate coverage once the domain itself has been selected for measurement. However, these procedures do not solve the more basic question of which domains, from the multitude of those that make up a curriculum area, should be included for measurement.

The premise of the B.C. program, like others of its kind, is that priority for measurement will be given to those learning outcomes judged to be basic or essential to all students. The important word here is "judged" and our experience with provincial representative panels of professional and concerned lay people has borne out the fact that consensus as to what is meant by "basic" and "essential" is not easily obtained. A major cause for this difficulty may be the fact that definitions of these terms, as we noted earlier, have tended to change over time. Nevertheless, if assessment is to proceed, this fundamental question of what to measure must be resolved.

Of related concern is the tendency to measure that which is easiest to be measured. Those of us involved in large-scale measurement activities give frequent lip service to the need for eclectic information collection procedures, and at the same time we continue to rely almost exclusively on paper and pencil instruments in multiple-choice format. Indeed, as Nix (1979) recently observed, large-scale assessments tend to shy away

from hard to measure curriculum areas such as oral communication in favor of areas such as reading and mathematics which we find amenable to traditional paper and pencil techniques. The obvious danger in this rationale is that some people — the public in particular — may come to view these areas as less important simply because we do not report on them in our assessment activities. Conscientious efforts to develop and validate other data-gathering procedures must continue to be made if large-scale assessments are to reflect adequately the scope and complexity of today's curriculum.

Assessments conducted in British Columbia in the areas of social studies (Aoki et al., 1977), written expression (Conry et al., 1978) and physical education (Carre et al., 1979) have incorporated into their design some innovative procedures for measuring pupil performance. As a spin-off from some of these assessments, the measurement procedures have led to the production of resource materials packaged for teacher use following the assessment. We believe the support of this type of development to be an important and integral role of large-scale assessment programs if the state of the art is to be advanced.

Similarly, advances must also be made in our ability to document the link between the processes and the products of education. This information must be available if rational decisions regarding the improvement of provincial programs are to be made. While recognizing the difficulties in establishing causal relationships in a system as multidimensional as a province's educational program we at the same time feel strongly that finding ways to improve the degree to which process and product instrumentation are complementary is critical to the generation of intelligent recommendations for change.

A final measurement issue is related to complaints of over-testing. Instructional time is finite, and as decision-makers at varying levels within the educational system launch their own assessment and evaluation activities, this finite time is increasingly eroded by testing activities. The problem for a provincial assessment program becomes one of ensuring adequacy of content coverage to achieve desired content validity and measurement precision, while at the same time not overly extending the administration time. This problem is somewhat compounded in B.C. by the fact that matrix sampling can only be used to a limited extent when individual reports are to be produced for each district, some of which are very small reporting units (less than 50 students in some grades). As well, it would be desirable to develop possible solutions to eliminate unnecessary duplication of effort at the various levels of the educational system. Latent trait procedures as potential solutions to these problems have begun to appear with increasing frequency in the research literature (Bauer et al., 1979; Hambleton & Cook, 1977; Hill, 1979). The British Columbia Ministry of Education is completing the first stage in the development of a calibrated mathematics item bank with the intent that

it will serve as the basis for development of a variety of instruments, including multiple forms of survey assessment tests. But while latent trait models may serve to provide viable solutions to some of the more thorny measurement problems of assessment programs, further investigations are in order before these approaches can be seen as the panacea to our measurement ills.

Interpretation of Assessment Information

Assuming that the decision has been made as to what to measure, the problem becomes one of answering the question, "How well have students performed?" Notwithstanding the critical review given to the process of standard-setting by Glass (1978), it remains that if assessment data are to be used for purposes other than description, judgments regarding what are acceptable and unacceptable levels of performance, whether for an individual or a group, must be made. The proliferation of publications and conference symposia on this topic attests to the interest in this question, both on theoretical and practical grounds. As the review by Hambleton et al. (1979) makes clear, there are a variety of procedures for establishing standards of performance ranging from purely judgmental to largely empirical. What is now needed is comparative examination of these various procedures and an exhaustive review of the question of bias.

A concern that we have frequently voiced in B.C. is over the confusion between norms and standards. For example, we have found in both our mathematics and science assessments that B.C. students outperformed students in the United States on a number of test items common to the National Assessment of Educational Progress (NAEP). If we were to use the traditional approach to the interpretation of test results, we would be saying that our system is doing very well because we outperformed some other system. But, if as educators we are now accountable for the education of all students, one can't be satisfied with the fact that, according to our data, large numbers of young adults are entering the job market lacking important skills — notwithstanding the fact that we are doing better or worse than time past, or better or worse than some other jurisdiction. Experience dealing with interpretative panels, involving both educators and members of the public, for six assessments has repeatedly demonstrated an initial skepticism to any method of interpretation that is other than normative.

Akin to this problem is resistance to recognizing that difficult subjective decisions must still be made if assessment results are to be translated into priorities for action. Establishing the link between assessment data and proposed recommendations for follow-up is implicitly assuming a causal relationship between process and product. In British Columbia we collect a variety of process and product information in the belief that this larger data base will ultimately lead to better recommendations. However,

because these links cannot be clearly established, recommendations for change, out of necessity, are largely based on a combination of professional judgment and conventional wisdom.

The fundamental question concerning interpretation is how far an assessment should go. Should such a program get involved in actually formulating recommendations and establishing priorities for follow-up action? Or is the NAEP model to be preferred wherein performance is described, compared with previous performance levels, but left to the reader to interpret for its educational implications? It is our position that the former is the appropriate stance for an assessment program launched by a provincial educational agency. At the same time, it behooves the measurement specialists associated with such a program to inform the managers and policy-makers of the very real limitations of the data and of other sources of relevant and perhaps competing information. This last point leads logically into the third area of concern, that surrounding communication of results.

Communication of Assessment Information

Perhaps the single greatest need in the area of assessment is for reporting of results in easily understandable terms to the potential user without at the same time distorting the results through over-simplification. The need is for brief summaries that do not distort the results and that direct the reader to other sources of more detailed information. In our experience, the tendency is to overburden the policy-maker with excessive information to the degree that major points are lost in the deluge of detail. At the same time we sense a real need to provide our readers with increasing sophistication in the appropriate interpretation of assessment data; we see this as a major need for the future.

An important purpose of the majority of large-scale assessment programs is accountability, that is, to inform the public of the current status of public education. As such, assessment reports are in the public domain, and in the course of events quickly find their way into the public media — often on the front page. What makes the best copy from the vast array of assessment information may not always be what the authors of the report, in fact, believed to be most crucial or relevant. What is topical, no matter how trivial, tends to receive greatest attention.

A reverse problem can arise when other less credible but persuasively presented messages are the catalyst for decisions which may be completely the reverse of those indicated by the assessment results. Assessment reports, particularly if they tend to be dry, scholarly affairs, are hard pressed to compete with influential groups or individuals who make "good copy."

A variety of publics are concerned with the outcomes of education. We believe it to be a need to identify these groups early in any assessment

activity to ensure that communication lines are opened before misconceptions arise through lack of information. The alternative is that information which is useful and potentially valuable is collected but is either ignored or misinterpreted by those who ultimately make the educational decisions.

Utilization of Assessment Information

The assessment program in British Columbia was designed on the assumption that decisions made about education are best made if based upon up-to-date information on what and how students are learning. This statement assumes first that the information collected is that which the decision-maker needs, and second that the decision-maker is prepared to use it in his deliberations. Experience has taught us that these assumptions are frequently not warranted. Conscious efforts must be made on the part of the assessment program staff early in the planning of an assessment to identify questions which decision-makers wish to have addressed. This is more often than not a difficult task, as Nix (1979) has also recently observed:

> Many managers still do not see measurement as an essential function in their repertoire of managerial competencies. Too often they make decisions with limited information, do not participate in assessment and evaluation design, and view measurement as a nuisance to be tolerated at best, or a source of adverse criticism at worst.... Without collaborative effort among managers, curriculum experts, and psychometrists in generating relevant assessment or evaluation questions, measurement runs the hazard of missing its mark with respect to providing information which can and will be used by managers and policy makers for the improvement of instruction. (p. 22)

Of related concern is the potential for abuse of assessment information. For example, assessment data can just as easily be used to support poor as good decisions. To illustrate, a small district in B.C. was concerned about the performance of its Grade 4 students on five questions dealing with dictionary skills. Its response was to allocate almost the entire resource book budget in the following year to the purchase of dictionaries for all Grade 4 students. Questions such as "What else was not purchased that year because of this allocation?" and "Will the purchase of materials result in a change in the problem area?" come immediately to mind.

A problem closely related to this one is the tendency of assessment data to put the educational system into a reactive, short-term mode, rather than the reverse effect that it is intended to have. Indeed, assessment by its very nature is designed to allow an educational system to be more rational in its long-range planning and establishment of priorities. When headlines of less than satisfactory assessment results appear on the news stand, there is the tendency for the public and their elected representatives to want immediate action, of the sort referred to in the example

above. Because resources are finite, this short-term response may be at the sacrifice of other long-range programs and activities which, in the full scope of education, are of greater priority.

No amount of assessment data collection or report distribution will ever be guaranteed of having a positive influence on the educational system or of eliminating misuses and abuses. Are there steps that can be taken to minimize the incidence of misinterpretation and maximize positive follow-up? Shepard (1978) has observed that assessment programs are ways by which the public is made aware of the status of public education. These same publics must also be made aware of the limitations inherent in any one set of information. Accountability demands cannot be adequately fulfilled by assessment results which measure only a small part of what goes on in schools.

A FINAL NOTE

Over the past several years, ministries of education across Canada have been facing increased pressure to establish comprehensive accountability schemes. Response to these pressures has been complicated by the fact that over a relatively short period of time in history, society appears to have embraced the principle of universal education, and is now demanding for all students literacy standards which at one time were expected of only a small segment of the student population. The growing trend toward the development of provincial assessment programs — one form of response to the demand for greater accountability — presents educators with a dilemma. On the one hand there would appear to be general agreement over the claim that a provincial educational system needs to document the progress of student learning, but on the other hand it is becoming clear that the problems associated with provincial assessment programs are not, at the present time, easy to resolve. Given that decisions can be made with or without province-wide assessment data, however, the choice is whether or not educators wish to use this vehicle to participate in or influence the decision-making process. It is our position that the answer can only be "yes" if responsible management of education is to ensue.

NOTES

[1] "Assessment" is a word which appears to be adopted by the education community to distinguish large-scale measurement programs from more traditional testing programs aimed at measuring individual student achievement (Shepard, 1978).

[2] Grade 12 scholarship examinations are still administered in British Columbia.

REFERENCES

Aoki, T. T., Langford, C., Williams, D. M., & Wilson, D. C. *Goals and situational studies: British Columbia social studies assessment* (Vol. 1). Victoria: B.C. Ministry of Education, 1977.

Bauer, E. A., Shawski, E. J., & Veitch, W. R. *How minimal is minimal?* Paper presented at the annual convention of the NCME, San Francisco, 1970.

British Columbia Ministry of Education, Learning Assessment Branch. *Guidelines for interpretation: District assessment results, 1976.* Victoria: B.C. Ministry of Education, 1976.

British Columbia Ministry of Education, Learning Assessment Branch. *Guidelines for interpretation: District assessment results, 1977.* Victoria: B.C. Ministry of Education, 1977.

British Columbia Ministry of Education, Science and Technology, Learning Assessment Branch. *Guidelines for interpretation: District assessment results, 1978.* Victoria: B.C. Ministry of Education, 1978.

Carre, F. A., Mosher, R. E., Shutz, R. W., Thomson, P., Farenholtz, E., & Bullis, L. *The B.C. assessment of physical education.* Victoria: B.C. Ministry of Education, Science and Technology, 1979.

Conry, R., Muir, W., Mussio, J., Rogers, T., & Wideen, M. *Assessment planning: British Columbia assessment program.* Report submitted to the Joint Committee on Evaluation, July 1975. Victoria: B.C. Department of Education, 1975.

Conry, R., & Rodgers, D. *The British Columbia assessment of written expression: Summary report.* Victoria: B.C. Ministry of Education, 1978.

Evanechko, P., Armstrong, R., Dey, J., Johnson, T., & Quorn, K. *Language B.C.: Summary report.* Victoria: B.C. Department of Education, 1976.

Glass, G. V. Standards and criteria. *Journal of Educational Measurement,* 1978, *15,* 237-261.

Greer, R. N. *Some factors related to reading and mathematics achievement: District level analyses.* Victoria: B.C. Ministry of Education, 1978.

Hambleton, R. K., & Cook, L. L. Latent trait models and their use in the analysis of educational data. *Journal of Educational Measurement,* 1977, *14,* 75-96.

Hambleton, R. K., & Eignor, D. A practitioner's guide to criterion-referenced test development, validation and test score usage. *Laboratory of Psychometric and Evaluative Research Report No. 70.* Amherst: School of Education, University of Massachusetts, 1978.

Hambleton, R. K., Powell, S., & Eignor, D. R. *Issues and methods for standard-setting.* Paper presented at the annual meeting of the National Council for Measurement in Education, San Francisco, April 1979.

Hill, R. K. *Use of the Rasch model to solve data problems encountered by the California Assessment Program.* Paper presented at the annual meeting of the National Council for Measurement in Education, San Francisco, 1979.

Hobbs, E. D., Boldt, W. B., Erickson, G. L., Quelch, T. P., & Seiben, G. A. *British Columbia science assessment: Summary report.* Victoria: B.C. Ministry of Education, 1978.

Millman, J. *Strategies for constructing criterion-referenced assessment instruments.* Paper presented at the Eighth Annual Conference on Large-Scale Assessment, Boulder, Colorado, June 1978.

Nanaimo School District. *Report of the Nanaimo school district interpretation of results of the Reading Assessment, 1977.* Report submitted to the B.C. Ministry of Education, March 1978.

Nix, C. W. *Educational measurement needs in the 1980's with emphasis upon statewide assessment.* Paper presented at the annual meeting of the National Council of Measurement in Education, San Francisco, April 1979.

Popham, W. J. *Criterion-referenced measurement.* Englewood Cliffs, N.J., Prentice-Hall, 1978.

Putman, J. H., & Weir, J. M. *Survey of the school system.* Victoria: King's Printer, 1925.

Resnick, D. P., & Resnick, L. B. The nature of literacy: An historical exploration. *Harvard Educational Review.* 1977, *47*(3), 370-378.

Robitaille, D. F., & Sherrill, J. M. *The British Columbia mathematics assessment: Summary report.* Victoria: B.C. Ministry of Education, 1977.

Schofield, R. S. The measurement of literacy in pre-industrial England. In J. R. Grady (Ed.), *Literacy in traditional society.* Cambridge: Cambridge University Press, 1968.

Shepard, L. *Purposes of assessment.* Paper presented at the annual NAEP conference, Boulder, Colorado, 1978.

Traub, R., Wolfe, C., Evans, P., Russell, H. H., & Wayne, K. *Secondary-postsecondary interface project II: Nature of students* (2 vols.). Toronto: Ontario Ministry of Education and Ministry of Colleges and Universities, 1976.

Victoria School District. *Report of Greater Victoria School District follow-up to the reading and mathematics assessments, 1977.* Report submitted to the B.C. Ministry of Education, Science and Technology, Victoria, 1979.

Jerry J. Mussio is Director of the Curriculum Development Branch and R. Nancy Greer is Director of the Learning Assessment Branch of the British Columbia Ministry of Education, Victoria, British Columbia v8v 2M4.