DOCUMENT RESUME

ED 323 262 TM 015 509

AUTHOR McEwen, Nelly; Chow, Hau
TITLE Issues in Implementing Indicator Systems.
PUB DATE Jun 90
NOTE 32p.; Paper presented at the Annual Meeting of the Canadian Educational Researchers' Association (Victoria, British Columbia, Canada, June 5, 1990).
PUB TYPE Reports - Evaluative/Feasibility (142) -- Speeches/Conference Papers (150)

EDRS PRICE MF01/PC02 Plus Postage.
DESCRIPTORS Academic Achievement; *Educational Quality; Elementary Secondary Education; Foreign Countries; *International Programs; Outcomes of Education; Pilot Projects; *Program Implementation; *Systems Development
IDENTIFIERS Canada School Achievement Indicators Project; *Educational Indicators; Educational Quality Indicators Project (Alberta); International Educational Indicators Project; *Quality Indicators

ABSTRACT

A conceptualization of educational indicator systems is presented, and their guiding principles and criteria are outlined. Three developmental projects are described that represent the indicator systems that promise to provide an explanatory framework for education. An indicator system should be perceived as an integrated whole in which the various aspects function as a system. Multiple indicators are needed to reflect the complexities of education. Any indicator system should be based on the goals, priorities, and expectations of its originating group, whether local, provincial, national, or international. The three examples that illustrate a broad perspective on indicator systems that attempt to embed student outcomes in the context of the educational system are: (1) the International Educational Indicators Project focusing on Australia, the United States, France, Austria, the Netherlands, and Canada; (2) the Council of Ministers of Education, Canada School Achievement Indicators Project; and (3) the Educational Quality Indicators Project of Alberta (Canada). These examples illustrate the importance of involving the stakeholders in the establishment of an effective and integrated system of indicators. A 79-item list of references is included. (SLD)

TM

Issues in Implementing Indicator Systems

Nelly McEwen and Hau Chow
Planning and Policy Secretariat
Alberta Education

Paper presented at the annual meeting of the Canadian Educational Researchers' Association, Victoria, June 5, 1990

Issues in Implementing Indicator Systems

This paper presents a conceptualization of educational indicator systems and outlines their guiding principles and criteria. It presents three developmental projects which represent the new generation of indicator systems that promise to provide an explanatory framework for education. These are particularly relevant to the Canadian context because they explore local, provincial, national and international perspectives of education. The paper subsequently addresses the major issues related to the implementation of indicator systems, based largely on the difficulties which have manifested themselves in the American context where indicators have been corrupted in attempts to produce short-term gains to satisfy annual accountability demands in an essentially long-term endeavor. It classifies the issues into three categories – educational and social, administrative and technical, and political and financial – and discusses their implications. Positive and negative effects are identified and strategies proposed for minimizing the negative impacts. The final section proposes strategies for resolving the issues.

Indicator Systems

This section describes the principles and criteria for indicator systems. The underlying premise is that education is a complex enterprise which cannot be described by means of any single component. To examine student outcomes in isolation, without taking into account the context and operation of schooling, denigrates a large body of research on the relationships among achievement and other variables, and may perpetuate poor pedagogical policies and practices. This discussion takes a systems approach which by definition is an integrated whole, rather than a number of discrete parts. A new generation of indicator systems applying these principles is illustrated by three current developmental projects.

Conceptualization

According to Webster's dictionary (1983), a system is a: "regularly interacting or interdependent group of items forming a unified whole, ... having a common purpose, [with] a form of social, economic or political organization or practice [whose] manner of classifying, symbolizing or schematizing [t . a] harmonious arrangement or pattern" (p. 1199). There are many different evaluation models, but a commonly used one based on a systems perspective is the context, input, process, product (CIPP) model. Stufflebeam and Webster (1988) described the objectives, methods, and relation to decision making in the change process of each of the four types of evaluation based on the CIPP model. They defined two types of functions for evaluation: formative (decision making) and summative (accountability) and provided guidelines for deciding when to

undertake which type. Their guiding principle is that educational institutions exist fundamentally to foster human growth and development. Therefore, assessments of students' needs must take into account all relevant developmental variables, and assessments of institutions must take into account all components and enterprises that bear on meeting the needs of their students. To guide institutions to identify and assess the performance needs of students, seven categories of student development were articulated: intellectual, emotional, physical and recreational, moral, aesthetic and cultural, vocational, and social (p. 584). They concluded their chapter as follows:

> Evaluation is a vital function in the offering of educational services. No system can achieve its potential and maintain a high level of service if it does not constantly assess its performance and modify its practices accordingly. This is as true in the individual classroom as it is in the office of the superintendent or the principal. Those in positions of leadership can help their systems adopt, implement, and use a sound and wide-reaching program of evaluation ... Possible actions pertinent to the improvement of evaluation ... [may] be manifested in increased and improved evaluation practice in school districts and schools (p. 598).

The current emphasis on educational reform has led to greater accountability and an increase in monitoring and evaluation of schools and systems. In order to determine what type of information to collect, many organizations are developing and implementing indicator systems. Indicators can paint a broad picture of the conditions of education and stimulate thinking about potentially effective policies (Shavelson, 1988, p. 6). While indicators cannot describe a system completely, they can inform planning, policy and decision making (McEwen and Zatko, 1989, p. 7). The goals, or intended benefits, of implementing indicator systems are to improve education, to provide a mechanism for accountability, and to assess the efficiency and effectiveness of the educational enterprise. An educational indicator system should provide:

- information that is feasible to gather, valid and useful for policy decisions;
- logical clusters of indicators to increase the system's interpretative power;
- points of reference such as a previous value, the value of a comparison group or some socially determined standard;
- measures of the ubiquitous features of schooling that are: enduring, easily understood, feasibly measured and generally accepted as valid and reliable statistics; and
- a reporting system that is accurate and timely (Oakes, 1986; Selden, 1988).

The existing American accountability indicator systems were precipitated by the release of *A Nation at Risk* and other reports in 1983 which identified a dearth of information about schools and schooling (Selden, 1988). According to Finn (1988), the reform movement in the United States arose from widespread anxiety that deteriorating educational performance was endangering the spirit, economic vitality, technological prowess and perhaps even the national security of the

United States, while eroding the ability of individuals to have a happy, rewarding, safe and virtuous life (p. 116). The reform movement insisted on tangible evidence of results, most commonly in the form of rising scores on various tests and other indicators of cognitive learning. This resulted in the high degree of homogenization or standardization of education, with a commensurate narrowing of the range of choice given to students and the breadth of discretion given to teachers. According to the Office of Educational Research and Improvement (OERI, 1988), 46 states have accountability systems of which 35 are state systems, two are local, and nine are a mixture of both state and local systems (pp. 28-29). The major types of indicators reported are: student achievement on standardized testing, attendance and completion of school. While 45 states collect background data, only 21 reported using these data for interpreting performance indicators. Large-scale testing programs form the core of the state accountability systems and often drive the local response. Questions about the current status and use of standardized tests are being raised increasingly by researchers, the media and the public. The most common criticisms include the inability of existing tests to measure the full range of achievement; the limited overlap between the substance of standardized tests and the skills and content stressed in textbooks; a general narrowing of the curriculum as a result of the tests; emphasis on basic skills; and the confusing messages transmitted to policy makers and the public as a result of the long intervals between revising norms (OERI, 1988; Brandt, 1989; Meisels, 1989; Koretz, 1989).

Change or reform cannot occur without information. The role of an indicator system should be one of providing information for specific purposes, such as for planning, setting policy and making decisions. Once information has been collected and made available, the distinction between its use for improvement or accountability becomes blurred. The use of public information cannot be controlled. Therefore, what information is reported to whom in what format is a critical issue.

An indicator system is NOT synonymous with a student testing or assessment program. Much of the criticism levelled against the existing American accountability systems is due to their use as programs of student assessment without regard for the consequences or effects that their introduction produces. An indicator system should be a much broader, more comprehensive undertaking which attempts to provide a more balanced picture of the operation and outcomes of schooling, particularly if the indicators focus on a broader range of desired outcomes. By focusing on student learning, and interpreting the findings in terms of educational conditions, better decisions may result from the enhanced information (McEwen, 1990, p. 13). Student outcomes are but one, albeit vital, component in this larger picture of educational conditions.

The New Generation

A new generation of indicator systems is currently under development. These new systems have the potential to reduce the negative impact of the American accountability systems which focus almost exclusively on student outcomes and have resulted in the negative consequences for students and educational institutions. Three examples illustrate a new and broader perspective on indicator systems which attempt to embed student outcomes in the context of the educational system provided through schooling.

The Organization for Economic Cooperation and Development (OECD) has completed its 18-month exploratory phase and is currently entering the developmental phase of the International Educational Indicators (INES) Project which will conclude December 1991 (OECD, 1989a). Originally INES was divided into five networks, each of which was the responsibility of a particular country. These were: 1. enrolment, educational career paths and school leavers at various stages (Australia); 2. education outcomes (United States); 3. the functioning of schools (France); 4. assessing costs and resources (Austria); 5. attitudes and expectations (Netherlands). Canada participated in networks 1 and 2. The next phase of INES will finalize the conceptual and analytical framework for the production of indicators and delineate an indicator system.

This international project adopted a systematic scheme as a framework for the development and interpretation of indicator systems: environment, resources, processes and effects (Nuttall, 1990). An understanding of the effects or outcomes of education must be informed by the educational processes employed and resources (fiscal and human) deployed, against the background of contextual factors in the environment of schools or of education systems. Six principles guide the work of this project:

1. Indicators are diagnostic and suggestive of alternative actions, rather than judgemental;
2. the implicit model underlying a set of indicators must be made explicit and acknowledged;
3. the criteria for the selection of indicators must be made clear and related to the underlying model;
4. individual indicators should be valid, reliable and useful;
5. comparisons must be done fairly in a variety of ways (eg, with like groups, with self over time, and using dispersions and differences between subgroups as well as averages); and
6. the various consumers of information must be educated about its use (p. 10).

The INES project is important because of its international scope and broad intents. It is exploring the major conditions of education which need to be examined in concert with desired outcomes if true educational improvement is to occur.

Another promising indicator project is the Council of Ministers of Education, Canada (CMEC) School Achievement Indicators Project to develop national indicators for education. All provinces and territories are participating. It consists of three interrelated components: I. participation, retention and graduation rates; II. expectations of and satisfaction with the education system; and III. student achievement of literacy and numeracy of 13 and 16-year-old students. While this project is less ambitious than INES, it will provide a Canadian information base that will enable provinces to assess the performance of their school programs in comparison with Canada-wide standards (CMEC, 1988). This project can make an important contribution to understanding Canadian educational results. There is no federal department of education since education is a provincial responsibility, so national information about education is released every five years through the census. The dearth of comparative educational studies in Canada makes it difficult to gauge the performance of the provinces and territories. Provincial results essentially exist in a vacuum. An exception is the Second International Science Study (of the International Association for the Evaluation of Educational Achievement [IEA]) which was the largest and most comprehensive Canadian curriculum study ever conducted (Connelly, Crocker and Kass, 1985; 1989). More than 20,000 students participated in this first cross-Canada survey of science achievement at grades 5, 9 and 12/13. The two volumes documenting the Canadian results provide descriptive data on Canadian curriculum policy, science classroom practice and student achievement. The CMEC indicators project will contribute ongoing information about Canadian literacy and numeracy. More studies of this type are needed if Canadians are to understand the consequences of their educational endeavors.

The third example is Alberta's Educational Quality Indicators (EQI) initiative. This is a collaborative endeavor between Alberta Education and twelve school jurisdictions to develop indicator systems to measure the success of the educational enterprise in the province. This three-year initiative has three components: 1. indicator systems (meeting three criteria: an interpretative framework to describe variation among students and schools; student outcomes related to the educational enterprise; and points of reference for comparing results); 2. methods (to collect, analyze and interpret the information); and 3. reporting and dissemination (to inform diverse audiences of the results). A four-dimensional model of education was developed to guide the direction of this initiative. It consists of partners (schooling, family and society), conditions (context, inputs and processes), student outcomes (cognitive, affective and behavioral) and time (grades 3, 6, 9 and 12) (McEwen and Zatko, 1989).

Each participating school jurisdiction is currently developing and will implement a local indicator system which represents its educational goals and priorities and reflects the local community's expectations. The information generated from the projects will assist Alberta Education to develop provincial indicators. The interpretation and recommended directions of the

local indicator projects, together with other provincial initiatives, will provide a solid foundation for the implementation of an efficient and effective information system which measures the success of the educational enterprise in the province (McEwen, 1990).

Indicator systems – whether local, provincial, national or international – are interdependent because they reflect the goals and priorities of their originators. The common goal is an educated citizenry to deal with the increasing complexity of life in an interdependent world. A shared priority is the development of the next generation as efficiently and effectively as possible given competing demands for limited resources. No community can operate in isolation today. Education must prepare students to take their place as responsible citizens. It is both a social and economic imperative that students be prepared to contribute to their society (community, province, country) because we all live in the global village. For any community to survive, it must provide economic opportunities or else its residents leave. The current depopulation of rural Canada testifies to the inability of the agricultural sector to provide a living for many people who are flocking to urban areas, where chances of earning a livelihood are greater given the larger population and the range of employment opportunities this provides. Kanter (1989) discussed how the turbulent, fast-paced environment of global competition and constant innovation in technology and corporate structure has dramatically changed our lives (p. 11). The impending ecological disaster is another example of the interdependence of nations. "Saving the planet" cannot succeed unless all countries strive to solve global problems. Everyone has a part to play in making the world a better place for all of us. Education is a social mechanism for promoting shared values and societal goals.

The next sections present the issues for each of the three major areas and discuss their implications. Both positive and negative effects are identified. In many cases issues result from opposing ideas of what constitutes educational excellence, equity, effectiveness and efficiency. These originate in cyclical swings of conservative and liberal values. The resolution of issues occurs when compromise takes place.

Educational and Social Issues

This section examines the educational and social issues of implementing indicator systems. Addressed are: purposes and priorities, curriculum and assessment, impact on students, structure of schooling, and context and values.

Purposes and Priorities

In the 1980s educational reforms emanated from several sources, each having a particular perspective (Futrell, 1989; Cuban, 1990). Reform was motivated by political or economic proponents who view education as serving the national interest, and by educational and democratic

proponents who view schools as enabling students to reach their potential. These opposing viewpoints reinforce the issue of the function of schooling, be it for instrumental and/or intrinsic purposes. All educational authorities have defined goals of schooling and/or education. These are often approved by the legislature and provide the direction for the instruction of children in a province. Most goals tend to be of three major types: academic, societal and personal. Programs and activities are planned, taught and evaluated on the basis of these goals for students. While most goal statements tend to be broad and encompass major areas for development, existing provincial assessment programs tend to focus on a narrow range of these goals. Subjects which are not assessed are perceived to be less important and can fall victim to restraint.

All educational authorities (local, provincial, or national) establish priorities in order to allocate scarce resources. The 1988 Alberta School Act established five principles to ensure the fundamental purpose of education – to ensure that students learn; these principles are: access to quality education, equity, flexibility, responsiveness and accountability (Alberta Education, 1987, p. 4). The present Minister of Education, Jim Dinning, added three priorities to guide the direction of education in Alberta – excellence, equity and efficiency. These are similar to those proposed by Boyer (1989) who identified five priorities for the systemic rather than symbolic push for excellence: goals; equality; teachers; school leadership; and accountability. These are but two examples of how priorities set the agenda for education.

Curriculum and Assessment

The curriculum is usually divided into core or essential programs and complementary or personal interest programs. At the elementary level, there is usually a clearly defined group of programs that all students pursue; at the secondary level, the core consists of English, social studies, mathematics and science – and it is expected that all students learn the rudiments of these academic subjects. Complementary programs often include the fine and practical arts which are viewed as being of interest to some students, but not necessary for all. The current curricular reforms across Canada have defined core and complementary somewhat differently which has led to considerable variation among the provinces in what is defined as a basic set of expectations for students (for example, Alberta [1985] and British Columbia [1988]). In their documentation of the Canadian component of the Second International Science Study, Connelly, Crocker and Kass (1985) described the diversity in education in Canada: in policy, organization, styles of teaching, curricular emphases and development, language of instruction, and in rights to instructional services (p. 311). This diversity introduces problems when contemplating the implementation of national indicators.

Probably the most visible and most common indicator of student outcome or performance has been some type of test score. Assessment and its impact on the curriculum, the learning process and the educational system as a whole has been the focus of several recent special issues:

Educational Leadership (April 1989); *Phi Delta Kappan* (May 1989); *Educational Researcher* (December 1989); and the *Alberta Journal of Educational Research* (March 1990). The impact of the use of standardized tests on the curriculum is related to the use of these test scores as accountability tools. The critique of the use of tests in this context is mostly negative (eg, Koretz, 1989; Neill and Medina, 1989; Wiggins, 1989). The pressure to raise test scores has resulted in a disruption of the learning process because the priority becomes raising short-term test scores for the class rather than ensuring that students learn. This can result in practices such as teaching to the test and a general narrowing of the curriculum, where the curricular content is dictated by the content of the test (Shepard, 1989). Although test developers attempt to use current curricula to guide test construction, the resultant test may not be responsive to changes in curricula over time and differential emphasis across different systems. Another factor that contributes to the disassociation between test scores and learning is the fact that test scores are usually not given much weight by the teachers. McLean (1985) in his survey of student evaluation in Canada found that although teachers may explicitly allot time to prepare students for the tests, the results are usually not used (see also OERI, 1988). The scores are not used directly to calculate students' marks and most schools do not really study the results closely. The mismatch between the test content and what is taught in the classroom plus the emphasis on lower-order thinking skills may be contributing factors. The end result is that while the district/government and the public place a lot of emphasis (testing is usually funded) and credibility (schools are held accountable) on test scores, this is not translated into instructional practice because teachers see this type of testing as an unnecessary intrusion into the classroom.

In questioning the need to examine outcomes, Nagy (1990) concluded that: testing technology distorts the curriculum; political realities interfere with doing the job properly; people want to see "hard" numbers rather than anecdotal reports, even though these hide the real data and threaten people. Moreover, evidence suggests that the intended, translated and achieved curriculum are different (eg, Connelly, Crocker and Kass, 1989). Hargreaves (1989) recommended the following strategies for reforming curriculum and assessment: decentralization of curriculum development; administrative support for a collaborative teacher culture; mandatory guidelines requiring a broad and balanced curriculum and reinforced through the power of inspection; and a revamped assessment system designed to provide teachers with improved feedback about their pupils and their progress as a basis for curricular renewal (p. 170).

Impact on Students

Nagy, Traub and MacRury (1986) discussed in detail the impact of proposed province-wide testing in Ontario of selected high school courses. Some of the effects on students are: test anxiety, fear of failure, increased number of students entering the general stream and increased number of students

dropping out of the system. The impacts on the teaching-learning processes are: intrusion of the exams into the classroom; less emphasis on discovery-oriented learning; and more emphasis on strategies to raise test scores. The impacts on institutional policies are: the use of scores as accountability tools for schools; and use of scores as an admission criterion to post-secondary institutions.

The implementation of an indicator system has to deal with the issue of individual differences in students, the most prominent of which is differences in ability levels. The use of tracking or streaming – students in different programs/courses based on prior achievement – has been the most common method. Although tracking may be beneficial for some groups of students in the high tracks, it is an ineffective and unfair way of dealing with individual differences because of its damaging effects on students in the low tracks (Oakes, 1987). Oakes identified some of the effects of tracking as: students in the low tracks do not receive the remediation that might enable them to move to a higher track; students in vocational tracks do not receive the adequate or appropriate training to compete in the work force; the achievement gap between students in high versus low tracks widens over time; decisions made in the lower grades have long-lasting effects on placements in different tracks; for students in the lower tracks, the process tends to foster lowered self-esteem, lowered aspirations and negative attitudes toward school. Oakes examined some of the school and societal contexts for the continued use of tracking despite its negative impacts. She suggested that future research on whether tracking works should consider how, for whom, and toward what ends. In an Ontario study of the relevance of education and the issue of dropouts, Radwanski (1987) recommended the elimination of streaming as well.

Structure of Schooling

Tyler (1984) said that "education ... is a social enterprise seeking to help persons acquire understanding, skills, attitudes, interests and appreciation" (p. 29). The structure of schooling needs to be examined if educational improvement is to occur. Some of the models of education which contribute to our understanding of the relationships between student learning and development and schooling follow. Carroll (1963) described the degree of learning as a function of time spent (determined by opportunity to learn and perseverance) over time needed (determined by aptitude, ability to understand instruction and the quality of instruction). Hymel (1988) expanded Carroll's model to include four additional theories: Bloom's mastery learning model (student characteristics, instruction and learning outcomes); Edmonds'/Brookover's effective schools correlates (instructional leadership and quality, instructional focus, school climate, teacher expectations and behaviors, program evaluation via standardized measures); Spady's premises of outcome-based schooling (capability of students to excel, success/self-concept/learning linkage, alterable instruction so as to improve learning, strategies to maximize learning conditions); and

Hunter's mastery teaching concepts (content, learner style and behavior, teacher behavior). Since 1975 Walberg and his colleagues have been trying to develop a comprehensive framework for the analysis of productivity and to test it in a variety of classroom studies (Walberg, 1984, pp. 20-21). Nine factors require optimization to increase student (affective, behavioral and cognitive) learning: student aptitude (ability, development, motivation), instruction (quantity and quality) and environment (home, classroom, peer group, mass media). Oakes' (1986) comprehensive model of the educational system consists of three parts: inputs (fiscal and other resources, teacher quality, student background); processes (quality of: school, curriculum, teaching, instruction) and outputs (achievement, participation and dropouts, attitudes and aspirations). Shavelson and his associates (1987) developed a framework for a possible indicator system which piggybacks onto ongoing data collection through the National Assessment of Educational Progress. Their piggyback indicator system looks at inputs (resources, teacher characteristics, student characteristics), processes (school and classroom characteristics, and instruction); and student outcomes (achievement, participation, attitudes).

Coleman (1986) discussed the good school district in the context of three major descriptions of instructionally effective schools: Goodlad's *A Place Called School* (1984), Lightfoot's *The Good High School* (1983), and Rutter et al's *Fifteen Thousand Hours* (1979). Common features of the three studies included: the focus on 'culture' as opposed to resources; the consequent emphasis on "thick description"; and the holistic concern that syndromes, configurations, and patterns are important, rather than single variables such as class size. The good school has a dual focus: it pursues concurrently academic (focusing on learning, accountability and change) and nurturance (focusing on caring, collegiality and community) purposes (p. 90). He concluded his discussion by suggesting that good school districts have characteristic norms and practices, labelled elements, which have classroom, school and district-level consequences. The linkage among these elements is coordination (p. 95).

In their review of effective schools, Purkey and Smith (1983) found some commonality of findings: strong leadership by the principal and other staff; high expectations by staff for student achievement; a clear set of goals and an emphasis for the school; an effective schoolwide staff training program; and a system for monitoring student progress (p. 435). The authors proposed the following nine organization-structure variables as important: school-site management; instructional leadership; staff stability; curriculum articulation and organization; schoolwide staff development; parental involvement and support; schoolwide recognition of academic success; maximized learning time; and district support (pp. 443-445). In terms of a future agenda, they recommended longitudinal studies in a variety of schools that track school and student performance over time; a fuller investigation of the process by which schools increase, decrease or maintain effectiveness; identifying information about the procedures employed, the obstacles encountered, and the intended

and unintended results obtained; and finally, identification of the intermediate steps of "goals specification and problem diagnosis" (pp. 447-448).

In her review of schooling research, Oakes (1989) identified three global school conditions that act as enablers for high-quality teaching and learning: access to knowledge, press for achievement, and professional teaching conditions. Mediating factors include school resources, alterable organizational policies and structures, and school culture. School resources are necessary but insufficient for quality education; the way resources are allocated is also important. Some alterable policies and structures are: allocation of time (for instruction); curricular emphasis; grouping practices (tracking based on abilities); extra support for low achieving students; and amount and kind of parental involvement (especially at the elementary level). Aspects of school culture that are conducive to quality are: commitment to student learning; primacy of teaching; collaborative staff planning/sharing and teamwork; and opportunities for program improvement and professional renewal. Global school conditions function synergistically – they act as enablers rather than as causes of quality teaching and learning.

The OECD international report of schools and quality (1989b) concluded that "student motivation and achievement are profoundly affected by the distinctive culture or ethos that is to be found in each school" (p. 141). The report identified ten characteristics that appear to play a role in determining school outcomes. Some are similar to those identified by Oakes (such as collaborative planning; staff development; and parental involvement); others include the commitment to clearly and commonly identified norms and goals; leadership in initiating and maintaining improvement; staff stability; maximum use of learning time and the active and substantial support of the responsible education authority (p. 140).

The preceding discussion on the structure of schooling illustrates why indicator systems need to incorporate these alterable conditions. Information derived through these indicators can suggest possible intervention strategies for improving education.

Context and Values

Context variables, such as home and school environment, are important in the conceptualization of student learning. In his nine-factor model, Walberg (1984) identified home environment as one of the causal factors in student learning. In his synthesis of over 3,000 studies of student learning, Walberg concluded that the "alterable curriculum of the home" is twice as predictive of academic learning as family socioeconomic status. This "curriculum" refers to "informed parent-child conversations about school and everyday events; encouragement and discussion of leisure reading...expressions of affection and interest in the child's academic and other progress as a person" (p. 25). Fraser, Walberg, Welch, and Hattie (1987) in identifying the salient facets of student learning, stated that "the pupil is part of a home, school, and classroom and is subject to

the influences from peers, teachers, parents, and the media" (p. 193). In other words, the amount of learning any individual child can demonstrate is partially determined by context factors outside the control of the educational system. Although a child's previous achievement level remains one of the most important variables in determining student outcome, that level is influenced by the child's home environment.

The fabric of Canadian society has changed enormously over the past 25 years (Health and Welfare Canada, 1989). These changes have placed tremendous pressure on families as they try to prepare their children for an increasingly complex and fast-paced society. The traditional, two-parent family with the mother staying at home has been largely replaced by dual-income families which have steadily increased since the 1960s, from 25% to over 60%. This has indirect effects such as an increased demand for quality child care. The pressures faced by families in the 1990s are more than just financial. The number of one-parent families has increased significantly over the past 25 years. From 1961 to 1986, one-parent families increased from 8.4% of all families to 12.7% (Statistics Canada, 1961, 1986). The rates for Alberta were 5.6% and 11.8% respectively (Alberta Education, 1989). This increase is in part due to the dramatic rise in the divorce rate: the rate for Albertans was 78.0 per 100,000 people in 1961; in 1986 it was 394.3 (Alberta Education, 1989).

The ethnocultural structure of Canada is also changing due to shifts in immigration patterns. "Immigration is far more diverse now than it was a quarter of a century ago in terms of the national, linguistic, religious and racial backgrounds of immigrants" (Health and Welfare Canada, 1989, p. 32). Immigrants from the United Kingdom alone accounted for over 25% of all immigration in 1966; they accounted for less than 10% in 1986. The shifting patterns have a direct and immediate effect on the educational system. The growth of the number of non-English/non-French speaking school-aged children places pressures on the educational system for increased resources (staff, time, and money) for language training. The increasing multicultural aspect of Canadian society also demands that the educational system address these complexities.

Education does not and cannot exist in a vacuum. Changes in the economic, social, and cultural aspects of society impact on the child's social, emotional, and cognitive development. The issue becomes how schooling can accommodate differences in the children's learning potential as a result of the constantly changing society. It is important to understand the context in which schooling takes place. Techniques to control for background influences are required. Nagy, Drost, and Banfield (1985), in an investigation of community isolation and its influence on educational achievement, used path analysis to control for economic and parental effects. They found a substantial contribution for some of the isolation variables in accounting for achievement differences. While background variables are not amenable to manipulation by educators, they have a great impact on achievement effects and need to be examined.

The search for evidence of differences in the effectiveness of schools has become an important theme in educational research. Madaus, Kellaghan, Rakow, and King (1979) questioned the use of standardized tests as measures for comparing the quality of different schools. In their study of secondary schools in Ireland, they found that by using curriculum-based tests instead of standardized tests, they were able to demonstrate the effectiveness of schools above and beyond the child's background (which was the major conclusion of Coleman *et al* (1966) that "schools bring little influence to bear on a child's achievement that is independent of his background and social context" [p. 325]). Madaus *et al* used a different methodological approac' for dealing with variations from the family, the individual and the classroom. Instead of entering single variables as predictors in regression analyses, variables were grouped into blocks and then used as predictors in stepwise regression analyses. One of their conclusions was that school factors are the main contributors to explained variance in curriculum-based tests while individual factors are the main contributors for standardized tests.

Cuban (1990) conceptualized the struggle over reform as a conflict over values. In a historical perspective on recurring reform, he described the changes that have occurred in schooling over the last century and a half of pendulum swings: universal access to education; introduction and spread of the graded school; changes in governance; certificated staff; expanded curricula and extracurricular activities; and reduced class sizes. When conservative values were prevalent (in the '20s, '50s and '80s), schools were concerned with producing individuals who could compete: hence, high academic standards, orderliness, efficiency and productivity were prized in schools. When liberal values dominated (in the early 1900s, '30s and '60s), concerns for minorities, the poor and outsiders prompted school reforms to broaden student access to programs, linked schools to work in the community, and reduced academic achievement gaps between groups of students. Each political turn of the cycle left a residue within schools when the rhythm shifted. Cuban offered two alternative explanations (which he termed speculative) to the political cycle of liberal-conservative, of public versus private interest – the political and institutional perspectives.

Many issues which reappear are value conflicts. When economic, social and demographic changes create social turmoil, public opinion shifts. Differences, as they become transformed by the media and political coalitions into pressure on schools to change, are dilemmas that require political negotiation and compromises among policy makers and interest groups. This is what occurs in the larger society. There are no solutions, only political trade-offs. Cuban stated that people turn to schools in times of social turmoil because dominant social groups want public schools to work on national ills, rather than risking major dislocations in society, by directly addressing major social problems. People have a shared, enduring belief that schools promote social mobility, creating national harmony, and building solid citizens. Reformers transform schools because they are instruments to express and maintain the prevailing ideologies of a society.

While reforms that strengthen prevailing beliefs get implemented, some reforms are for display, not fundamental change. Cuban's second alternative explanation for why reforms fail is the institutional perspective. Schools are different from nonpublic organizations in that they are tax supported and under lay governance. The combined political and institutional perspectives help to explain why districts in different parts of a province(state), region and nation resemble one another in structures, roles and operations. District operations are tightly coupled in meeting legal requirements, avoiding conflicts of interest, and spending funds. However, the daily delivery of instruction is virtually decoupled from administration and policy making because teacher support is required to provide instruction. Therefore, reforms return but seldom substantially alter the regularities of schooling.

Administrative and Technical Issues

There are many administrative and technical issues involved in implementing indicator systems. Included in this discussion are: strategic considerations, resources, system design, reliability and validity, and data bases.

Strategic Considerations

When an indicator system is contemplated, four strategic questions dominate the discussion:

1. How can the information be collected and coordinated?
2. How can it be used effectively and efficiently?
3. Who is responsible for the management of the information?
4. Who is responsible for maintaining the currency of the information?

Currently different jurisdictions (local, provincial, national) collect a large volume of information (such as annual education statistics, census data, achievement and ability test results, surveys and polls). The problem is that the information is collected, analyzed and reported in different ways so that it is difficult to compare and integrate the information from the different sources. The implementation of an indicator system may mean developing a way to incorporate existing information or modifying collection procedures to provide integrated information. For example, Fetler (1989) used the California State Department of Education data base to investigate factors that contribute to the dropout rate; these included: total enrolment, enrolment in academic courses, achievement scores, and index of poverty. The study illustrates the conceptualization of a model for predicting dropout rates utilizing existing data bases. The most difficult and crucial issues are the conceptualization of the problem and the selection of the appropriate variables.

When information is collected by more than one level of government and/or school jurisdiction, there must be a way to aggregate and disaggregate the information. Different levels of the educational system require information at different levels of specificity. Information collected at the individual or classroom level by the teacher likely needs to be aggregated to at least the school level to be useful to the district. Nagy (1990) distinguished between differing information needs at the local and provincial levels. What may be important to local authorities intent on improvement at the micro level, might not be the same as that considered by provincial authorities intent on improvement at the macro level. Summary data serving political and accountability purposes may not provide diagnostic information which the school can use for prescribing action. Bock and Mislevy (1988) recommended the duplex design as a way to accommodate differing information needs and to tailor information reports to the specific needs of different users such as students, schools, districts, government, and the media. McEwen and Zatko (1989) also described differing provincial and local information needs. Provincial governments are concerned with achieving a standard which is comparable to national and international counterparts. School jurisdictions are concerned with comparing and interpreting results for their own students in different schools as well as comparing their results to other jurisdictions of comparable size, location and economic base. The two perspectives require some common indicators for provincial purposes and others for local or regional needs.

An important strategic consideration is coordination. People responsible for indicator systems need to have considerable technical expertise to appreciate the complexities of research design and implementation (such as appropriate measurement methodologies, reliability, validity, etc). They also need highly developed administrative skills to deal with the management, coordination and dissemination of information to appropriate audiences. Moreover, they need persuasive skills to negotiate the desires and expectations of various stakeholders (administrators, teachers, parents, the public, others), and to be able to communicate clearly and effectively with the different interested parties (government, school personnel, research community, the public). People who possess a combination of all these skills are rare, and yet essential to ensure that the indicator system can capitalize on the various levels of expertise provided by the different groups.

Resources

Implementing an indicator system requires the commitment of considerable resources – information, staff, money, and time. A discussion of each resource follows. McDonnell and Elmore (1987) described three types of information to assist in shaping the choice of policy and/or indicators: political (what is preferred by other policy makers, organized interest groups and constituents); strategic (information about the target, its capacity to implement, and its probable response to various policy instruments); and analytical (information about the technical requirements of various

instruments and which ones are likely to work under different conditions). The last two types of information are particularly important for policy areas such as education, in which control is fragmented among different actors and levels of government (p. 25). Oakes (1989) concurred with the importance of information about resources, policies, organizational structures and processes that characterize schools. In recognition of the importance of informed discussion, both OECD and Alberta Education are providing technical documents to support their respective indicator projects.

Communication of information is a vital component in successful educational change. Connelly, Crocker and Kass (1988) categorized communicative devices into two categories: progress and methods. Devices to ensure and measure progress include proposals, reports, position papers, and the use of consultants and advisory committees; communication methods include meetings and computer communications. They found the most important functions of meetings were to solidify interpersonal relations and to highlight deadlines and progress to allow decisions to be made on priorities. Records of meetings served as a chronology of the project and were useful for task orientation and interpretation. Electronic communication was an efficient and less expensive method than the telephone to maintain contact among the research teams (pp. 446-447).

Staff commitment and expertise are also important ingredients for the successful implementation of an indicator system. Without shared expectations and the support of the community in the determination of the goals, priorities and expectations for its students, a local school jurisdiction may not address the questions its constituents want answered. Lack of involvement in and development of commitment to such an undertaking may lead to attempts to thwart the system. Odden and Marsh (1988) recommended collaboration and initial involvement of key players to build commitment. They also recommended using cross-role teams, training and assistance, and continuing leadership support and pressure as ways to keep people interested in and committed to projects.

Financial assistance is also necessary for the implementation of any innovation. In times of economic restraint, however, it is often a question of reallocating existing resources rather than finding new sources of revenue. One way to deal with the financial issue is to allocate resources in kind: for example, releasing staff members from current obligations to undertake new tasks, reassigning priorities for existing staff and financial commitments, and eliminating redundant or unproductive activities.

The importance of time cannot be underestimated. Assessing students takes considerable time; indicators must be selected judiciously so that the minimum amount of instructional time is required for measuring student outcomes. Sampling strategies such as those proposed by Maguire and Rogers (1989) and McLean (1987) can assist to minimize the amount of time any individual

student needs to be involved. Information about the conditions (context, inputs and processes) may not disrupt instructional time, but exacts its toll in terms of collecting, analyzing and reporting. Deploying more staff to complete these tasks can reduce elapsed time, but may result in less accuracy because of the greater number of people working on different parts. Competing demands of accuracy and timeliness become a major issue.

Kanter (1989) summarized the issue of resources by suggesting capitalizing on existing strengths and tackling opportunities while living within constraints. The challenge of doing more with less is the ultimate balancing act (p. 372).

System Design

The implementation of an indicator system presents many technical issues such as sampling, data collection, generalizability, reliability, and validity. Two research projects exemplify the differences between a centralized and a decentralized mode of operation. Frymier *et al* (1989) detailed some of the techniques used in a collaborative research project of students at risk with 276 schools in 87 Phi Delta Kappa (PDK) chapters. Data were collected from over 22,000 students and over 9,000 teachers. There were four questions common to all chapters plus specific questions for individual chapters. The basic strategy was to provide a detailed instructional manual (140-page document) for each participating chapter. The manual served as the common denominator for the different sites by standardizing data collection procedures and measures, thereby increasing uniformity across the different study sites. Reliability of the data collection was increased by having a small group of trained coders code the data from all sites. Analyses were also centralized, using the same statistical procedures and programs. This study illustrates a practical methodology for research with organizations that have both centralized and decentralized features.

The PDK study can be contrasted to the approach and the resulting issues identified by the Canadian research team involved in the Second International Science Study (Connelly, Crocker and Kass, 1988). Data were collected from over 20,000 grades 5, 9, and 12/13 students and their teachers. Canada was divided into three regions: western (the territories and the western provinces); central (Ontario), and eastern (Atlantic provinces). Within each region, the project was coordinated by a principal research team at a university. In effect, the project was akin to three separate but simultaneous projects because each of the coordinators had to conduct his/her portion within the jurisdiction of his/her university. Each team solicited funding and research support on its own. Some of the problems encountered related to differences in research capabilities/support of the investigators and the affiliated university. Unlike the PDK study, there was no central organizing and coordinating structure. Effective communication, ranging from planning meetings to transmitting data and reports, became a major issue.

Reliability and Validity

For the establishment of any system of indicators, the mechanisms or tools for measuring the indicators must be reliable and valid. Although reliability of the instruments used to measure the indicators is always important, validity is paramount. That is, one should not steer away from potential indicators that are meaningful (high validity) but difficult to measure (low reliability). The current discussions on the need for more authentic student assessment of overall achievement aim to develop measures of highly meaningful skills, such as critical thinking, in more reliable ways (Frederiken and Collins, 1989; Norris, 1989; Wiggins, 1989). This is contrasted to the current reliance on easy-to-score, multiple-choice questions that emphasize lower-order thinking skills. McLean (1987) noted that many of the well-constructed, psychometrically sound tests have strong validity, for purposes of sorting and ordering students, but have low pedagogical validity since the tests do not suggest ways to improve learning. Ideal measurements should reflect school learning (the traditional notion of validity) and suggest improvements (pedagogical validity). McLean called pedagogical validity the soul of assessment, and accountability its body or visible manifestation (p. 4).

Messick's (1988) distinction between evidential and consequential validity of an assessment instrument is also relevant for indicator systems. He argued that "the process of test use inevitably places test scores in both a theoretical context [evidential basis] of implied relevance and utility and a value context [consequential basis] of implied means and ends" (p. 41). Test developers and users have to be critical of these aspects of test use and interpretation when evaluating its validity. Some of the social consequences of (achievement) test use are teaching to the test and narrowing of the curriculum. When an indicator system is implemented, it is often the social consequences that are most difficult to deal with. As it is the test administrator's ethical responsibility to ensure that tests are used appropriately and scores interpreted within the context of local values, the organization responsible for the implementation of an indicator system bears the same sets of responsibilities. There are clearly defined guidelines regarding the use and dissemination of test results (for example, American Psychological Association, 1986; Code of Fair Testing Practices in Education, 1988; ETS Standards for Quality and Fairness, 1987; Standards for Educational and Psychological Testing, 1985). This is not the case with indicators such as demographics, student-teacher ratios, salaries, etc. Their nontechnical nature may lead to misuse and misinterpretation because of their deceptive simplicity. The actual meaning or interpretation of these types of indicators may be extremely complex.

Data Bases

An essential requirement for developing and implementing an indicator system is the establishment of a data base. Aside from issues of reliability and validity, one of the criteria for judging a data base is its currency. Currency of a data base should not be confused with the use of a data base to track yearly or periodic trends. The issue of currency for standardized tests relates to judging individual performance against the normative sample provided by the test developer. Koretz (1988) cited the lack of currency of the norms of standardized tests as one of the primary reasons for the Lake Wobegon effect (where many states performed above average). Since renorming a standardized test is an expensive and time-consuming process, the test user is forced to make trade-offs between an instrument that has excellent psychometric properties but dated norms and another one that may be weaker psychometrically but has more up-to-date norms. One can supplement dated norms by maintaining local norms at the district or provincial level. Another solution is to combine normative information from the developer and to set a standard/criterion for specific schools/districts. Currency is also a concern for nontest score data bases. Once a commitment is made to establish a data base, it is important to ensure that there are established procedures and mechanisms for updating the information.

Another aspect of establishing and maintaining a data base is the consistency and comparability of the information collected. These issues become more complex as the number of reporting units increases and the number of instruments proliferates. For example, there are several instruments for measuring self-esteem. Similarly, a school system may note that nutrition is an important indicator of physical health, but individual schools may use different proxy measures of nutrition. How can one achieve comparability between schools/systems when an indicator is being measured by different instruments? Comparability is also an issue when measures evolve and change over time. Test developers in the cognitive domain have developed techniques (such as vertical and horizontal scaling) for dealing with comparability. Techniques to deal with comparability for instruments in the affective and behavioral domains have generally been less sophisticated (such as correlation coefficients). Frymier *et al* (1989) essentially dealt with this problem by insisting that individual PDK chapters use the same set of measures for the predetermined indicators/variables and allowing flexibility in others.

Political and Financial Issues

Political and financial issues are probably the most complex and difficult to resolve. Education is not independent of the political, social, economic and cultural forces interacting to shape the direction of a nation. Schooling is a powerful institution for developing the next generation, but its

support depends on the perceived value of its utility and effectiveness. "Government of the people, by the people, for the people" (Lincoln, 1863) at all levels – local, provincial and national – will determine what priority is assigned to that investment.

Governance

Centralizing and decentralizing authority in governing schools is an issue because many people are empowered to deliver educational services. During the '80s, centralizing authority gained support from policy makers who pursued school improvement through legislation. New reform proposals (Cuban, 1990), however, propose to decentralize decision making to schools based on the effective schools contention that the school is the unit of change and the corporate sector which delegates decision making to the site of production or delivery of services.

At issue is the question of responsibility and accountability for results. Since all members of the educational community contribute to the expected outcomes of the educational system, all must be involved, in some way, in the development and implementation of an indicator system so that they have a commitment to and investment in it. Top-down initiatives can be ignored because they do not have the support of the field; conversely, grassroots endeavors can be derailed because they cannot gain the interest of the people at the top. If education is a province's investment in its future, then all partners must contribute to its success: the provincial government must, in cooperation with its constituents, provide the direction and support for policies and educational programs which meet public expectations; school boards must manage the legal and financial operation of their schools, and accommodate the needs and priorities of their communities; schools must provide educational experiences for their particular client group of children; teachers must assume responsibility for the development of appropriate knowledge, skills and attitudes of their students; administrators need to ensure that the instructional staff has the necessary resources to fulfill their obligations; parents must work together with the school to enhance the opportunities for their children; and the public must support their local authority through taxation and moral support.

Education has dominated the government reform agenda for almost a decade. Government reform (Nagel, 1988) refers to changes in the structures and procedures of legislative, judicial and administrative institutions to make them more effective in achieving their purposes and more efficient in doing so with less time and expense (p. 10). The longer-term and broader effects of policy changes have benefitted society at large. Current trends in public policy include: higher goals for society in economic, social, political and science policy; increased benefits for both privileged and less privileged groups; the use of positive incentives for encouraging socially desired behavior; a more pragmatic, mixed approach in dividing responsibility between the public and private sectors; and the use of multicriteria decision making (pp. 13-14). This new expansionist philosophy emphasizes solutions to public policy problems that benefit all segments of society.

Hunter (1988) identified 13 multifaceted and controversial issues which have and will continue to dominate the public policy agenda. Among these are: education and training for a quality life and work force; refocusing on families and communities; and human and individual rights and values in practice. He suggested that a leadership challenge is to refocus on families and communities as key institutions in carrying out public policy while maintaining the rights and freedoms of individuals (p. 48). Among the strategies he proposed for managing change are recognizing: the global scope of issues; the long-term nature of issues; and the crossing of traditional boundaries (p. 55). The current leadership challenge is to avoid crises while making changes to reduce conflict and restore balance; that is, to inform and engage people in building their own future, by addressing the conflicts and imbalances in society and the economy and with other nations openly and honestly (p. 57).

Constraints and Capacity

Lam (1985) developed a two-dimensional matrix for measuring the constraints on s. hools. His conceptual framework can be extrapolated to encompass schooling (government and school jurisdictions). According to his framework, there are four domains which constrain schooling. The political includes policy and control; the economic includes funding and resources; the social includes values and demographics; and the cultural includes ethnic groups and language (pp. 372-374). Five environmental attributes – change, complexity, organization, routineness (certainty) and directness – can describe the external factors. This description can provide an understanding of the contingency approach as applied to the reality of the environment (p. 379). Zatko (1990) also commented on the difficulties in reforming education: the tension among top-down, collaborative and hybrid approaches to change; system inertia; the maintenance of sustained reform efforts; complexity and coordination; scope of reform; trust and collegiality; the differences between means and results; the desire for a blueprint versus a general direction; and the importance of communication.

Implementation of an indicator system needs to consider some of the forces in place that may either hinder or foster change. Firestone (1989) distinguished two aspects about the school system that influence possible actions toward reform: the will and capacity to use reform. The will to use reform is determined by the responses of the dominant coalition (usually consisting of the superintendent, the principal and other school staff) in the specific school district toward the proposed reform. The two critical dimensions of the group are: propensity for action and the perceived utility of the reform or policy. The most active use of reform results from a combination of a high propensity for action and a positive perception of utility of a particular reform or policy. Firestone identified three elements that are crucial for determining a district's capacity for reform: ease of personnel mobilization; necessary leadership qualities for the implementation of reform

(such as providing and selling a vision; obtaining resources; handling disturbances); and positive district-school linkages (such as conveying the attitude that the district is serious about reform; provision of targeted support; ownership of reform by teachers). In order for the implementation to be successful, an indicator system must be perceived as useful, the dominant stakeholder groups should have a high propensity for action; the organizational structure must allow for personnel mobilization; there is strong leadership; and there are positive district-school linkages.

Berry (1989) in describing the collaborative structure for outcome-based education, identified four interacting bases for transforming schools: vision (needs, beliefs, priorities), ownership (power, responsibility, decisions), capacity (role purpose, resources, operations), and support (organization and work structure). For successful change to occur, there needs to be a level of dissatisfaction with the existing order, a clear vision of what is desired, the identification of practical steps to effect change, and a belief that change is possible; these elements must be greater than the cost of change. Kanter (1989) suggested six criteria for forging successful partnerships and alliances over the long term:

- the relationship is important and gets adequate resources, management attention and sponsorship;
- longer-term investment tends to help equalize benefits over time;
- the partners are interdependent which helps keep power balanced;
- organizations are integrated so that contact and communication are managed;
- each is informed about the plans and directions of the other;
- the partnership is institutionalized through a framework of supporting mechanisms, from legal requirements to social ties to shared values, which make trust possible (p. 173).

Implementation Strategies

This section presents some strategies for minimizing negative impacts in the implementation of indicator systems. Odden and Marsh (1988), in developing a conceptual framework for studying education reform implementation research, stated that it should:

- integrate analysis of the content of the reform, the process of its implementation in the local setting, and its effects;
- focus on the influence of the reform on the overall local educational system as well as on the content, implementation process and more specific impacts;
- integrate a macro (provincial/state level) with a micro (district/school level) focus for analyzing the above issues;

- draw on the distinction between developmental and redistributive types of government programs;
- use recent research on the local change process and relate the results to the macro context to the content of the reform and to the outcomes at the local level; and
- identify several types of outcomes, including impacts on the individuals within local educational systems and impacts on the systems themselves (p. 45).

An indicator system is a tool for reform and improvement. Accountability is but one aspect of this endeavor. The information provided by an indicator system must be meaningful for all groups. Education is a shared responsibility between the provincial government, the family, schools, the academic community, and other institutions. Those making decisions about educational matters and those responsible for public funds must be held accountable for their decisions. School jurisdictions and provincial departments must use the results of their monitoring and evaluation processes to make appropriate adjustments to their educational systems (Zatko, 1990). Virtually all of the reports cited in the earlier sections identified the need to provide multiple perspectives on education in order to ensure that the indicator system will have the support of the key people in making it work. Therefore, the proposed strategy might be called the multiplier effect:

1. multiple goals of education – based on appropriate dimensions and domains of schooling;
2. multiple indicators of each goal – measured by multiple methods;
3. multiple levels of analysis: student, class, school, system, province, (and potentially) country, the world; and
4. multiple participants: government, administrators, teachers, academics, parents.

Indicator systems should be as flexible as possible to provide the above for planning, policy and decision making purposes. This implies that improvement is the goal, but that accountability must be demonstrated. There is no single appropriate indicator system, since each must be tailored to the needs, goals, priorities, and expectations of its originators. Thus while a provincial government might be concerned about determining the efficiency and effectiveness of its educational system based on credible criteria for assessing performance, the federal government might be interested in the comparison of Canadian students with their international counterparts on valued knowledge and skills.

Multiple goals are important because schooling is not a unidimensional activity. Most educational authorities have incorporated the development of student learning in cognitive, affective, behavioral and social domains into their programs of study. Yet little information exists about students' progress in developing desired concepts, skills, and attitudes. Measuring desired performance beyond the cognitive domain requires more sensitive instruments than either criterion

or norm-referenced testing approaches. Suggestions for developing pedagogical value in assessment were mentioned earlier (for example, McLean, 1987; Wiggins, 1989; Frederiken and Collins, 1989).

Moreover, multiple indicators of desired goals are important because each method of measurement has inherent philosophical and technical bases (Stufflebeam and Webster, 1988; House, 1983; Noblit and Eaker, 1988). If more than a single method is used to provide information about a particular goal, with its variations, more confidence will result in the findings. It is also easier to decide on improvement targets and strategies if the information comes from multiple sources. If parents, teachers and administrators all agree that student attendance requires improvement, based on a school's attendance records and anecdotal evidence, the consensus of what the problem is will result in better ways to ameliorate the situation. Applying these indicators to the CIPP model might suggest that low achievement and poor self-esteem (cognitive and affective outcomes) originating in an economically depressed inner city school (context) might be ameliorated by appropriate interventions (process) made by reallocating staff time and resources (inputs).

In addition to using multiple indicators for multiple goals, it is important to collect information so that multiple levels of analysis are possible. This has been referred to as micro versus macro analysis (Odden and Marsh, 1988; Cziko, 1989; Nagy, 1990) or as methods of aggregating data (Bock and Mislevy, 1988; McEwen and Zatko, 1989). Each of the various educational partners (students, parents, teachers, administrators, government, researchers, public) has different needs for information and each has a valuable perspective to contribute in interpreting results and suggesting improvement targets. Embedding multiple levels of analysis into the design of an indicator system will assist the different partners to share in the planning, setting policies and and suggesting instructional practices, and making appropriate decisions for their particular constituent groups. Analyzing local and provincial results is possible when common goals, curricula and assessment strategies are mandated. National and international comparisons introduce greater complexity due to differing educational philosophies, curricula, languages, organization, and so forth.

Finally, the key to the successful implementation of an indicator system is to involve the partners in the process. An indicator system will succeed only if it is viewed as valuable and important. While each group of partners has its own agenda for reform and improvement, they can work together to take action on their particular area of responsibility. Education, like other public policy areas, changes through negotiation and compromise among competing values and expectations. That this compromise should result on the basis of informed opinion, rather than through vested interest, is paramount. An effective indicator system, whose goals, analyses, and methods represent multiple perspectives on education, can make a major contribution to improving education and therefore, the economic and social prospects for the next generation.

Conclusion

The implementation of indicator systems is a complex activity fraught with potential pitfalls. Issues can be categorized as educational and social, administrative and technical, and political and financial. The negative impacts can be minimized if an indicator system is perceived as an integrated whole in which the various aspects function as a system. Multiple indicators are needed to reflect the complexities of schooling. There is no single recipe for an indicator system. Ideally it should be rooted in a conception of the important components of the operation and outcomes of schooling. The system should be based on the goals, priorities and expectations of its originating group – be it local, provincial, national or international. The more stakeholder groups that are involved, the greater the need for discussion and agreement as to the most salient and important aspects for monitoring. Successful implementation necessitates understanding the issues so that negative impacts can be minimized. Resolution often involves compromise among competing perceptions and values. The benefits of an effective integrated information system which provides a comprehensive picture of the broad aspects of schooling are worth the risk.

References

Alberta Education. (1985). *Secondary education in Alberta.* Edmonton: Government of Alberta.

Alberta Education. (1987). *Framework for a new School Act. Highlights.* Edmonton: Government of Alberta.

Alberta Education. (1989). *Education in Alberta: a compendium of educational and social statistics.* Edmonton: Planning and Policy Secretariat.

American Psychological Association. (1986). *Guidelines for computer-based tests and interpretations.* Washington, DC: Author.

Berry, S. (1990, February). *Collaborative structure for outcome-based education.* Paper presented at the annual conference on outcome-based education. Phoenix, AZ.

Bock, R. D., & Mislevy, R. J. (1988). Comprehensive educational assessment for the states: the duplex design. *Educational Evaluation and Policy Analysis, 10(2),* 89-105.

Boyer, E. L. (1989). School reform. A national strategy. *Vital Speeches of the Day.* October 1, 1989, 741-744.

Brandt, R. (1989). On misuse of testing: a conversation with George Madaus. *Educational Leadership, 46(7),* 26-30.

British Columbia Royal Commission on Education. (1988). *A legacy for learners: the report of the Royal Commission on Education 1988.* Victoria: Queen's Printer for British Columbia.

Carroll, J. B. (1963). A model of school learning. *Teachers College Record, 64(8),* 723-733.

Code of Fair Testing Practices in Education. (1988). Washington, DC: Joint Committee on Testing Practices.

Coleman, J. S., Campbell, E. Q., Hobson, C. J., McPartland, J., Mood, A. M., Weinfeld, F. D., & York, R. L. (1966). *Equality of educational opportunity.* Washington, DC: US Government Printing Office.

Coleman, P. (1986). The good school district: a critical examination of the adequacy of student achievement and per pupil expenditures as measures of school district effectiveness. *Journal of Education Finance, 12(1),* 71-96.

Connelly, F.M., Crocker, R.K., & Kass, H. (1985). *Science education in Canada. Volume 1: Policies, practices, & perceptions.* Toronto: OISE Press.

Connelly, F.M., Crocker, R.K., & Kass, H. (1989). *Science education in Canada. Volume 2: Achievement and its correlates.* Toronto: OISE Press.

Connelly, F.M., Crocker, R.K., & Kass, H. (1988). National curriculum research: problems in the metamethodology of studies concerned with state policy and local variation. *Comparative Education Review, 32(4),* 430-451.

Council of Ministers of Education, Canada [CMEC]. (1988). National education indicators program approved. *Liaison, 13(2),* 2.

Cuban, L. (1990). Reforming again, again, and again. *Educational Researcher, 19(1)*, 3-13.

Cziko, G. A. (1989). Unpredictability and indeterminism in human behavior: arguments and implications for educational research. *Educational Researcher, 18(3)*, 17-25.

ETS Standards for Quality and Fairness. (1987). Princeton, New Jersey: Educational Testing Service.

Fetler, M. (1989). School dropout rates, academic performance, size, and poverty: correlates of educational reform. *Educational Evaluation and Policy Analysis, 11(2)*, 109-116.

Finn, C. E., Jr. (1988). Education reform in the United States: trends and issues. *International Journal of Education Research, 12*, 115-121.

Firestone, W. A. (1989). Using reform: conceptualizing district initiative. *Educational Evaluation and Policy Analysis, 11(2)*, 151-164.

Frederiken, J. R., & Collins, A. (1989). A systems approach to educational testing. *Educational Researcher, 18(9)*, 27-32.

Fraser, B. J., Walberg, H. J., Welch, W. W., & Hattie, J. A. (1987). Syntheses of educational productivity research. *International Journal of Educational Research, 11(2)*, 147-252.

Frymier, J., Barber, L., Gansneder, B., & Robertson, N. (1989). Simultaneous replication: a technique for large-scale research. *Phi Delta Kappan, 71*, 228-231.

Futrell, M. H. (1989). Mission not accomplished: education reform in retrospect. *Phi Delta Kappan, 71*, 8-14.

Goodlad, J. I. (1984). *A place called school*. New York: McGraw-Hill.

Hargreaves, A. (1989). *Curriculum and assessment reform*. Toronto: OISE Press.

Health and Welfare Canada (1989). *Charting Canada's future: a report of the demographic review*. Ottawa: Minister of Supply and Services Canada.

House, E. R. (1983). Assumptions underlying evaluation models. In G. F. Madaus, M. Scriven, & D. L. Stufflebeam (Eds.), *Evaluation models: viewpoints on educational and human services evaluation*. Hingham, MA: Kluwer-Nijhoff Publishing.

Hunter, K. W. (1988). Facing the public policy issues in 1989. *Futures Research Quarterly, 4(3)*, 37 - 58.

Hymel, G. M. (1988, April). *A macromodel of effective, outcome-based mastery learning school variables: an expanded view*. Paper presented at the annual meeting of the American Educational Research Association, New Orleans.

Kanter, R. M. (1989). *When giants learn to dance: mastering the challenge of strategy, management, and careers in the 1990s*. Toronto: Simon and Schuster.

Koretz, D. (1988). Arriving in Lake Wobegon: are standardized tests exaggerating achievement and distorting instruction? *American Educator, 12(2)*, 8-15.

Koretz, D. (1989). The new national assessment: what it can and cannot do. *NEA Today, Special edition - Issues'89*, 23-37.

Lam, Y. L. J. (1985). Toward the construction of a school environment instrument: a conceptual framework. *Canadian Journal of Education, 10(4)*, 362-382.

Lightfoot, S. (1983). *The good high school.* New York: Basic Books.

Lincoln, A. (1863). *The Gettysburg Address.* November 19, 1863.

Madaus, G. F., Kellaghan, T., Rakow, E. A., & King, D. J. (1979). The sensitivity of measures of school effectiveness. *Harvard Educational Review, 49*, 207-230.

Maguire, T. O., & Rogers, W. T. (1989). Proposed solutions for nonrandomness in educational research. *Canadian Journal of Education, 14(2)*, 170-181.

McDonnell, L. M., & Elmore, R. F. (1987). *Alternative policy instruments.* Washington, DC: Center for Policy Research in Education.

McEwen, N. (1990, April). *Educational quality indicators: developing indicator systems in Alberta.* Paper presented at the annual meeting of the American Educational Research Association, Boston.

McEwen, N., & Zatko, G. (1989, June). *Educational quality indicators: framework for action.* Paper presented at the annual meeting of the Canadian Educational Researchers' Association, Quebec City.

McLean, L. D. (1985). *The craft of student evaluation in Canada.* Toronto: Canadian Education Association.

McLean, L. D. (1987, April). *Emerging with honour from a dilemma inherent in the validation of educational achievement measures.* Paper presented at the annual meeting of the American Educational Research Association, Washington, DC.

Meisels, S. (1989). High-stakes testing in kindergarten. *Educational Leadership, 46(7)*, 16-22.

Messick, S. (1988). The once and future issues of validity: assessing the meaning and consequences of measurement. In H. Wainer & H. I. Braun (Eds.), *Test validity* (pp. 33-45). Hillsdale, New Jersey: Lawrence Erlbaum Associates, Inc.

Nagel, S. S. (1988). Trends in public policy in America. *Futures Research Quarterly, 4(3)*, 5-16.

Nagy, P. (1990, May). *Must we examine outcomes?* Paper presented at the annual conference of the Canadian Evaluation Society, Toronto.

Nagy, P., Drost, D., & Banfield, H. (1985). A path analytic study of the influences of community isolation on student achievement. *The Alberta Journal of Educational Research, 31(3)*, 209-227.

Nagy, P., Traub, R. E., & MacRury, K. (1986). *Strategies for evaluating the impact of province-wide testing.* Toronto: Ontario Institute for Studies in Education.

Neill, D. M., & Medina, N. J. (1989). Standardized testing: harmful to educational health. *Phi Delta Kappan, 70(9)*, 688-697.

Noblit, G. W., & Eaker, D. J. (1988). Evaluation design as political strategies. In J. Hannaway & R. Crowson (Eds.), *The politics of reforming school administration: the 1988 Yearbook of the Politics of Education Association* (pp. 127-138). Philadelphia, PA: The Falmer Press.

Norris, S. P. (1989). Can we test validly for critical thinking? *Educational Researcher, 18(9)*, 21-26.

Nuttall, D. L. (1990, April). *The functions and limitations of international education indicators.* Paper presented at the annual meeting of the American Educational Research Association, Boston.

Oakes, J. (1986). *Educational indicators: a guide for policymakers.* New Brunswick, NJ: Centre for Policy Research in Education, Rutgers University.

Oakes, J. (1987). Tracking in secondary schools: a contextual perspective. *Educational Psychology, 22*, 129-153.

Oakes, J. (1989). What educational indicators? The case for assessing the school context. *Educational Evaluation and Policy Analysis, 11(2)*, 181-199.

Odden, A., & Marsh, D. (1988). State education reform implementation. In J. Hannaway & R. Crowson (Eds.), *The politics of reforming school administration: the 1988 Yearbook of the Politics of Education Association* (pp. 41-60). Philadelphia, PA: The Falmer Press.

Office of Educational Research and Improvement [OERI]. (1988). *Creating responsible and responsive accountability systems.* Washington, DC: US Office of Educational Research and Improvement.

Organisation for Economic Co-operation and Development [OECD]. (1989a). *Project on International Educational Indicators (INES Project): review of preliminary results and proposal for future work.* Paris: Author.

Organisation for Economic Co-operation and Development [OECD]. (1989b). *Schools and quality. An international report.* Paris: OECD Publications Service.

Purkey, S. C., & Smith, M. S. (1983). Effective schools: a review. *The Elementary School Journal, 83(4)*, 427-452.

Radwanski, G. (1987). *Ontario study of the relevance of education, and the issue of dropouts.* Toronto: Ontario Ministry of Education.

Rutter, M., Maughan, B., Mortimore, P., Ouston, J., & Smith, A. (1979). *Fifteen thousand hours: secondary schools and their effects on children.* Cambridge, MA: Harvard University Press.

Selden, R. W. (1988). Missing data: a progress report from the states. *Phi Delta Kappan, 69(7)*, 492-494.

Shavelson, R. J. (1988). Contributions of educational research to policy and practice: constructing, challenging, changing cognition. *Educational Researcher, 17(7)*, 4-11, 22.

Shavelson, R. J., McDonnell, L., Oakes, J., Carey, N., & Picus, L. (1987). *Indicator systems for monitoring mathematics and science education.* Santa Monica, CA: The RAND Corporation.

Shepard, L. A. (1989). Why we need better assessments. *Educational Leadership, 46(7)*, 4-9.

Standards for Educational and Psychological Testing. (1985). Washington, DC: American Psychological Association.

Statistics Canada (1961). *1961 Census of Canada*. Catalogue #93-516. Ottawa: Author.

Statistics Canada (1986). *1986 Census of Canada*. Catalogue #93-106. Ottawa: Author.

Stufflebeam, D., & Webster, W. J. (1988). Evaluation as an administrative function. In N. J. Boyan (Ed.), *Handbook of research on educational administration* (pp. 569-602). New York: Longman.

Tyler, R. W. (1984). A guide to educational trouble-shooting. *Educational Leadership, 41(8)*, 27-30.

Walberg, H. J. (1984). Improving the productivity of America's schools. *Educational Leadership, 41(8)*, 19-27.

Webster's Ninth New Collegiate Dictionary. (1983). Markham, Ontario: Thomas Allen & Son Limited.

Wiggins, G. (1989). A true test: toward more authentic and equitable assessment. *Phi Delta Kappan, 70(9)*, 703-713.

Zatko, G. (1990, April). *The Alberta strategy for educational reform: balancing inputs, processes, and results*. Paper presented at the annual meeting of the American Educational Research Association, Boston.