Measurement of the Full Scope of Learning

submitted to

The Pan-Canadian Education Research Agenda

R.A. Yackulic
B.W. Noonan
University of Saskatchewan
December 30, 1998

Table of Contents

Measurement of the Full Scope of Learning

Abstract

R.A. Yackulic
B.W. Noonan
University of Saskatchewan

One of the consequences of the past decade of educational reform in Canada is that there is a high level of agreement on the stated goals for education.

The Victoria Declaration endorsed by Council of Ministers of Education, Canada (CMEC) expressed two common fundamental beliefs: 1) that education is a lifelong learning process; and 2) that educational jurisdictions in Canada share many common educational goals (CMEC, 1993). One effect of these beliefs within specific school curricula results in a broader range of learning outcomes than traditionally acknowledged. A second effect of the broader range of goals is that educators must develop different ways to assess learning outcomes.

This paper examines these effects from four perspectives. First, the paper presents a review of the pan-Canadian consensus on the goals of education and a review of some of the existing provincial/territorial indicator programs. Second, the paper summarizes contemporary conceptions of the types of learning involved in the broad range of curriculum goals. Both cognitive learning and personal/social learning are considered. Third, assessment issues that evolve from a broader perspective of learning are identified. Fourth, the paper presents an overview of "opportunity to learn" as a way to enhance the interpretation of large scale testing programs and as a way to better explain the expanded conception of learning.

Finally, the paper presents a number of questions which may contribute to a pan-Canadian educational research agenda.

Introduction

One of the consequences of the past decade of educational reform in Canada is that there is a high level of agreement on the stated goals for education. An examination of provincial/territorial goals statements shows that there is a consensus that public education has a mandate for a broad range of learning outcomes. The stated goals for education include, but are not limited to, long-term outcomes such as developing critical and creative thinking, acquiring skills for productive citizenship, and commitment to lifelong learning. These goals typically are complemented by the more traditional intended learning outcomes related to skills and knowledge in mathematics, language, science, and history.

The Victoria Declaration endorsed by Council of Ministers of Education, Canada (CMEC) expressed two common fundamental beliefs: 1) that education is a lifelong learning process; and 2) that educational jurisdictions in Canada share many common educational goals (CMEC, 1993). One effect of these beliefs within specific school curricula results in a broader range of learning outcomes than traditionally acknowledged. A second effect of the broader range of goals is that educators must develop different ways to assess learning outcomes.

The purposes of this paper are as follows: 1) to review the consensus among provincial/territorial goals of education; 2) to consider these goals with respect to current perspectives of classroom learning and assessment; 3) to outline the potential of the concept of *opportunity to learn*; and 4) to identify questions for future research.

The consensus on goals for education

Intriguingly, the broad scope of current educational goals mirrors the range of goals from a previous era. Re-phrased in contemporary language, Thorndike's (1912) aims of education – happiness, utility, service, morality, perfectionism, natural development, knowledge, mental discipline, culture, and skill – would well capture the current goals of many educational jurisdictions in Canada.

The endurance of, and consensus for, the multi-faceted set of goals emphasizes that education is about more than "reading, 'riting, and 'rithmetic"; personal values and skills, thinking strategies, and skills and values required to contribute to and participate fully within society are included in the educational mandate.

The consensus on the goals of education is important in developing a pan-Canadian educational research agenda and will be an important source of research questions for the foreseeable future. As well, renewed public interest in benchmarks and standards augurs that the full scope of learning needs to be assessed with reference to societal expectations. One of the ways in which the goals have been evaluated, in terms of policy analysis, has been the development of a indicator programs across the country. *Education Indicators in Canada* (1996), *Education Indicators for Atlantic Canada* (1996), and the *Saskatchewan Indicators Report* (1998) are examples of the type of system monitoring currently in practice. Ontario Education established The Education Quality and Accountability Office (EQAO) specifically to address this mandate. The

indicators reported through these programs include demographic and social contexts such as participation rates, expenditures, equity programs, pupil-teacher ratios, and other system-level factors. Student achievement as an indicator is typically presented as student performance in basic skill subjects such as mathematics, language or science. There is little information on student achievement in other areas of the curriculum such as practical and applied arts education, technology, or social, personal, and life skills. Thus, in the context of the provincial/territorial goals statements, student performance tends to be represented by one dimension of the goals of public education, and a rather narrow scope of classroom learning.

Just as the indicator programs include a narrow interpretation of student learning so, too, do the large scale testing projects. The School Achievement Indicators Project (SAIP), the Third International Mathematics and Science Study (TIMSS), and the many provincial assessment programs measure similarly limited scopes of learning and intended curriculum outcomes. One of the reasons for this relatively narrow interpretation of student achievement is that a major purpose for indicator and large scale testing programs is accountability to stakeholders and the public. The background and rationale for this focus on accountability has been described by McEwen (1995), Hodgkinson (1995), Earl (1995), Maheu (1995), and Fagan (1995) in a special issue of the *Canadian Journal of Education.* Although these accountability programs are often described as contributing to educational improvement, it is not clear how the results of large-scale testing or reporting on system indicators results in improved system or student performance.

Assessing the full scope of learning places additional demands on assessment principles and practices that were developed to monitor more limited learning outcomes. Authentic assessment, curriculum-based assessment, forms of performance assessment and other emerging approaches attempt to sensitively capture this full scope of learning.

How does one operationalize the scope of learning in a way that better reflects the goals of education? It is necessary to expand the conceptions of school achievement to include other dimensions of educational goals: critical and creative thinking skills, technological literacy, and interpersonal skills. For example, over the past decade in the United States there has been a strong emphasis on various forms of character education and, more recently, a re-emphasis on the role of the school in developing personal and social talents (Kelly & Moon, 1998).

In summary, provincially and territorially set goals for education identify a full scope of learning as the mandate for schools. The intended learning outcomes for students are related not only to traditional areas of academic achievement but also to personological factors such as independence, responsibility, and self-efficacy; and to personal and social skills such as leadership, and cooperation. Provincial, regional, and national indicator projects have included student performance as an important indicator of educational ouput but with a narrow interpretation of student learning outcomes. Given that there is, by policy, a commitment to the full scope of learning, how do educators measure and set standards for the expanded view of intended learning outcomes? Addressing this question involves exploring current theories of learning, examining principles and practices of student assessment, and examining the relatively new concept of *opportunity to learn*.

What do we know about learning?

To better understand the implications for addressing a full scope of learning it is necessary to examine some of the current conceptions of learning that influence curriculum development and assessment practices. The notion of re-conceptualizing achievement is not new. Cole (1990), Broudy (1988, cited in Cole, 1990), Noonan (1996), and others have emphasized that our conceptions of achievement must be more related to the long term goals of the education system. This in turn demands that our beliefs and understandings about learning and about assessment need to be examined.

Resnick (1996), concluding a discussion of problems inherent in measuring complex cognitive outcomes, wrote:

> "We need ways of defining situations in terms of their cognitive demands and opportunities so that we can begin to develop a cognitive theory of accomplishment. A cognitive theory of accomplishment would explain how situation and person interact to produce a competent performance ... Pragmatic philosophy called for this kind of interactionist theory of cognition. Now so too does the practical demand for new forms of assessment (p.17)."

In similar vein, Bisanz and Bisanz (1994) emphasized that "assessments of processes and knowledge need to be closely tied to cognitive developmental theories (p.142)."

To do an adequate job of assessment, knowledge of what-is-being-assessed is essential. This is particularly true as we move from outcomes which are directly observable (e.g., focussing a microscope) to the many important outcomes which involve inherently unobservable internal cognitive and psychological processes.

Cognitive learning and cognitive outcomes

The behavioural-psychology focussed learning tomes of the '60s and early '70s have been replaced by a general theory[1] of cognition that is well-suited to examining educational learning outcomes. Philosophers such as Ryle (1949) have long distinguished among *knowing about* knowledge, *knowing how* knowledge and *reasoning*. Theories of learning have caught up.

The general theory of cognition evolves from Lewin's (1935) fields, Tolman's (1932) cognitive maps, Bruner, Goodnow, and Austin's (1956) categories, and Hanson's (1958) system of propositions. Cognitive learning includes the acquisition of concept knowledge, knowledge of the relations among concepts, and simple and complex mental processes. Anderson (1990) described three categories of cognitive learning outcomes: *declarative knowledge* (knowing about, concept knowledge), *procedural knowledge* (knowing how) and *thinking strategies/problem solving* (mental reasoning). His and others subsequent research has resulted in a rich,

---

[1] There is, of course, no single "general theory of cognition" that can be found in a text or an academic journal. Neither is there a single theory of mechanics or of health. However, in an applied profession, an eclectic best-statement of an amalgam of distinct theoretical positions may be preferable to a plethora of specific theories.

well-supported theory which illuminates the outcomes of many subject matter curricula.

Meaningful learning of declarative knowledge involves attaching new information within a learner's structure of prior knowledge, building a multitude of meaningful connections amongst the bits of knowledge in the structure, and organizing the knowledge structure to facilitate efficient access. According to the model, individual learners must generate their own knowledge structures; knowledge is not transmitted by teachers or books.

Procedural knowledge refers to thought-action sequences that are useful within a restricted knowledge domain. *Starting a car*, *finding the area of a triangle*, *metering a poem*, and *using a species name to locate an organism within a biological taxonomy* are examples of procedural knowledge learning outcomes. Learning procedural knowledge involves: 1) learning the procedure; 2) practising and self-monitoring the procedure step-by-step to develop proficiency; 3) proceduralizing (shifting from performing a sequence of steps [open the car door, sit in the driver's seat, close the door, ...] to a single performance [start the car]); and, in some instances, 4) automating the procedure such that it is performed without conscious control or monitoring. The student learns to perform the procedure, not to describe or list the steps in the procedure. Curriculum developers and instructional planners may represent the procedure in words; the learner need not. As for learning declarative knowledge, active learner involvement is essential and is encouraged by many contemporary instructional strategies.

Thinking strategies and problem solving refer to complex mental processes that are useful in a wider array of situations. *Reading*, *analysing*, *detecting logical inconsistencies*, *constructing an argument, identifying key relationships*, and *critical reasoning* are representative of thinking strategies and problem solving. There is much speculation about the nature of these mental processes and suggestions for encouraging the acquisition of these skills abound; alas, empirical research and theory concerning the learning and instruction of these outcomes are less developed (Greeno, 1989).

It is important to note several additional learning principles. Although cognitive learning outcomes are generally described verbally; declarative knowledge, procedural knowledge and thinking strategies need not correspond to words in an individual's vocabulary. That is, we can know ideas – such as the smell of fresh baked buns – but be unable to express that knowledge adequately in words. Or we can know how to perform a task on a computer, but become tongue-tied when attempting to explain the process to someone else. This abstract, non-verbal nature of learning outcomes is a challenge when monitoring achievement in classrooms where verbal expression is the token of commerce. Indeed, it is a continuing battle for to ensure that assessment strategies possess adequate construct validity.

Another principle is that cognitive learning is gradual and cumulative, rather than instantaneous and discrete. Bruner (1966) referred to this as spiral learning. Knowledge of a concept (and elaborations of links with other concepts) expands with each encounter with the concept: even a young child can possess accurate knowledge of a triangle, but that does not imply that her knowledge is as extensive as a mechanical engineer's. Similarly, procedural knowledge and thinking strategies are honed through practice, often at increased levels of complexity. At how many grades might the objective, "The student will write a short story," fit? This gradual and

cumulative nature of cognitive learning also constitutes a challenge for assessment. Although an instructional objective might be similar across several contexts or grades, the way in which it is assessed and the criteria used may vary considerably.

Yet another principle is that learning is not synonymous with performance. A learner may possess knowledge but be unable to express it in written (or some other specific) form. A learner may be capable of calculating the area of a triangle, but be unable to determine that that was the key to solving a particular word problem. Or a learner may be capable of writing a short story but decide not to write a short story on a particular topic on a particular test.

This general model of cognitive learning is consistent with many instructional strategies independently developed by effective teachers. Taba's inquiry method (Joyce & Weil, 1972), Wale & Stager's (1978) Guided Design, reciprocal teaching, case studies, even the Canadian military's aging DEER (demonstration, explanation, execution, repetition) focus on the learning processes highlighted by the model. It is increasingly common for textbooks used in pre-service teacher education to contain several chapters concerning implementing the model in math, reading and science (e.g., Gagne, Yekovich, & Yekovich, 1993).

Cognitive learning: manifestation and measurement

Knowledge of how learning occurs is helpful, of course, when designing assessments. Even more important, however, is knowledge of how a student changes as a result of learning. Since cognitive outcomes involve mental changes, it is critical to know how the mental events are manifest. Again, the need for a theory of achievement or a theory of performance is apparent.

Achievement of declarative knowledge results in a richly elaborated, organized knowledge base which allows a student to understand and explain ideas, interpret situations, form new relations among ideas, predict outcomes, and use the information to solve problems. Achievement of declarative knowledge traditionally has been inferred from responses to short answer, multiple choice, fill-in-the-blank, and similar items which require that the student write about or identify what she knows. If not well-constructed, these methods are susceptible to monitoring *rote learning* rather than *meaningful learning*. Since declarative knowledge of a concept involves personalized understandings far beyond verbal statement of facts, definitions or relations; many researchers have developed effective techniques for assessing concept knowledge and understanding in more sensitive ways. Concept maps, judgements of concept similarity, concept tree structures, and think-aloud interviews are among the methods with demonstrated effectiveness in assessing declarative knowledge. Alas, although these techniques seem to be acceptably valid and useful at the classroom level, the improvement over inferences from well-designed traditional methods is modest; their fussiness may preclude their use in large scale assessments.

Achievement of procedural knowledge results in a student's capability to perform a particular mental process. In most instances, these mental processes are subject-matter specific and are taught explicitly to students. Procedural knowledge traditionally has been inferred from student performance on tasks designed to evoke the particular mental process (e.g., if the mental process is to determine the area of a triangle, the student is given a triangle or dimensions of a triangle and

asked to compute its area). When carefully implemented, such straight-forward techniques are effective. However, validity concerns related to the following are common: the choice and design of the task, content representativeness, inappropriate emphasis on written responses, the particular scoring systems used, and matching the assessment to the target level of learning (step-by-step, proceduralized, or automated). The great enthusiasm for *performance assessment* and *authentic assessment* (see Bateson, 1994 for a critical review) which arose during the 1980s, focussed on the problem of *choice and design of the task*, sometimes at the expense of other validity-related issues. Alas, much of this work ignored the fact that procedural knowledge is inherently unobservable. As Resnick (1996) noted, "One thing that is clear now is that performance assessment cannot develop on solid ground without much more explicit theories of situated cognition than are now available (p. 17)."

Achievement of thinking strategies and problem solving skills also results in a student's capability to perform mental processes. These skills, however, are complex and powerful and generally are useful in a broad range of situations. Only a small subset of these skills are taught explicitly to students. Some of these skills (such as constructing an argument or writing an essay) have been common in classrooms for decades. Much is known about these skills and sets of general tasks have evolved that evoke the target mental processes. When assessing essay writing skills, for example, it is common to structure the assessment task to engage students in activation of prior knowledge, research, planning, initial writing, self-monitoring, and so forth. At least as importantly, scoring rubrics have been developed which focus on the specific features of the student's response that are related to the target mental process.

Perhaps spurred by the usefulness of mature rubrics for assessing writing, rubrics are being developed for many other complex mental skills. Because they focus on the product of a mental performance rather than the performance itself, rubrics do not resolve the need for improved theories of mental performance. They are helpful for some well-understood skills; they have been less useful for general, complex skills such as problem-solving, critical thinking, and decision making. These latter skills are not often assessed explicitly; until Resnick's concerns (quoted above) are addressed, their assessment will remain a black art.

Personal and social learning outcomes

Personal and social skills constitute another broad category of learning outcomes. These skills are captured by goal statements referring to independence, self-efficacy, lifelong-learning, citizenship and similar constructs. Infrequently addressed by specific school curricula, responsibility for these goals often has been left to families, churches and extra-curricular activities. Increasingly, however, these goals are being embedded as specific curriculum objectives. Kelly & Moon (1988) provided a framework for describing personal talents: a combination of affective (emotional) and connative processes. The affective components (also referred to as'emotional intelligence'; Goleman, 1995) include the ability to monitor one's feelings and to use information to guide one's thinking and action. The connative component of personal talent, "... helps individuals set and achieve goals, accomplish tasks, and persist in the face of

obstacles. (It) includes attention control, action control, and self-regulation strategies (Kelly & Moon, 1998; p. 744)."

Similarly, schools are expected to assist in the development of social talents, as essential life skills. Successful people at school or in the workplace tend to have high levels of social intelligence, that is, ability to adapt to many different social situations. The socially-talented person is one who is able to pursue both personal and group goals (p. 745).

Despite the apparent consensus that developing personal and social talents is a desirable educational goal, there is not agreement as to how these talents are learned, best taught, or assessed. In some cases, development of these talents is attempted through existing curricula (e.g., social studies or life style courses). In other cases, specific programs, such as *Lions-Quest Canada*[2] are used to teach personal and social skills. Perhaps more commonly, development of these skills is left to the vagaries of informal learning through general school experiences and extra-curricular programs.

Assessing personal and social skills is more problematic than describing their acquisition. First, there is disagreement as to whether such talents are "... a narrow domain specific ability or a broad ability that is demonstrated ... across many different types of social situations (Kelly & Moon, 1998; p. 745)." Second, many of these goals are phrased abstractly and lack the specificity needed for assessment. Third, there are no readily accessible instruments or commonly accepted strategies to assess personal and social talents.

It would be easy to dismiss these goals as beyond the reach of both curriculum intervention and assessment. Such an approach would be consistent with tradition. Unfortunately without specific attention to these goals, policies may evolve which inhibit their attainment. For example, many jurisdictions have safe school programs aimed at improving the safety of learning environments. In some instances, these programs use video surveillance cameras in areas where students socialize. If students conclude that use of the cameras transfers responsibility for acceptable behaviour from themselves to an external agent (the principal), development of self-discipline may be compromised. After all, the goal is for students to be self-disciplined and behave in a socially responsible fashion on their own, not just when the camera is on. Similarly, it may be that many of the learning experiences contributing to self-efficacy, initiative and perseverence are encountered through participation in the school's clubs and sports teams (see Childress, 1998, for further discussion of the many learning goals addressed by such activities). Curtailing these extra-curricular activities (for funding or personnel reasons) may compromise attainment of the personal and social goals.

Alas, without explicitly allocating responsibility for addressing these goals and monitoring their attainment, we will be unable to determine whether or not educational systems aid or inhibit development of these key outcomes. At the very least, strategies should be developed to assess if the goals are being addressed and if they are being attained. Given the abstractness of the learning outcomes, proxy measures may be needed. Participation rates in extra-curricular programs,

---

2 *Lions-Quest* is a life-skills education program published in Canada with the support of *Lions International*. It provides training in skills such as conflict resolution, service education, and healthy lifestyles.

incidence of violence, student participation rates in community service, school climate, discipline rates, and measures of student-teacher relationships are among the indicators that might be considered.

Some additional issues

A major premise of public education is “what little blighters learn in their desks will somehow be useful to them when they are not in their desks”: that is, transfer of learning is an implicit goal of education. A theory of learning and subsequent assessment of learning outcomes must address transfer of learning. Alas, research in this area is just developing. Mayer (1987) observed that transfer of declarative knowledge is facilitated by richly elaborated, well-organized knowledge structures that effectively incorporate the learner’s prior knowledge and personal experience; transfer of procedural knowledge and thinking strategies is facilitated by “over learning”, proceduralization and automaticity. For some time, it has been recognized that transfer is also facilitated when the learning situation corresponds closely to the situation in which the learning will ultimately be used. These and similar findings encourage the use of realistic situations during learning. They also support the use of similarly realistic contexts during assessment.

Critical to the discussion of transfer of learning is anticipation of the contexts to which the transfer might occur. This is an uncommon component in most curricula.

Problems inherent in predicting, teaching for, and assessing transfer of learning may be clarified as currently-popular theories of *situated cognition* (Greeno, 1989) are refined. Situated cognition is the notion that learning outcomes are products of the activity, context, and culture in which they were learned (Brown, 1988). Theories of situated cognition suggest that situation-specific knowledge is more powerful than knowledge intended for more general use. Cognitive apprenticeship (engaging students in ‘authentic’ learning activities in ‘authentic’ situations, much like a trade apprenticeship) is presented as a strategy for ensuring that cognitive learning is situated in contexts similar to those where it will ultimately be used. The implication of this is that transfer of learning is tightly context-bound; perhaps so should its assessment.

Bereiter (1997) was less enthusiastic about situated cognition. He noted that research has focussed on educational processes during learning and encouraged greater attention to learning outcomes. In similar vein, Anderson, Reder, and Simon (1997) described issues related to transfer of learning on which there seems to be agreement. The four areas include i) that it is possible to generalize from the classroom to the real world, ii) that knowledge can transfer across tasks, iii) that instruction can be in the abstract, and iv) that instruction need not take place only in complex social situations (p. 18).

Undoubtedly, there will be major implications for the assessment of learning outcomes as theories of transfer and situated cognition mature.

Assessing the full scope of learning

Expanding the range of learning outcomes calls for reconsideration of assessment practices.

For example, Masters and Mislevy (1991) and Mislevy, Yamamoto, and Anacker (1991) are among the many who have concluded that traditional methods of measuring achievement are inadequate to assess procedural knowledge or thinking strategies. Traditional methods are likely even more deficient when used to assess personal and social skills. Fortunately, many recent developments in assessment address the challege:

I. richer testing formats – simulations of decision situations have long been used to assess reasoning skills. The construct validity of these simulations may be enhanced when context information is presented using multi-media. BC Education and Saskatchewan Education have both used video clips effectively to present examinees with rich contexts. The Medical Council of Canada has developed computer-administered tests which simulate patient diagnoses (Blackmore, 1998). These approaches may be particularly useful for creating simulations for assessing social intelligence and similar outcomes.

II. growing consensus concerning indicator projects – Provincial and territorial indicator projects have established an assessment technology for basic accountability measures. For example, special issue of *Canadian Journal of Education* (20, 1, 1995) provided a review of the rationale, principles and practices of large-scale accountability testing in British Columbia, Alberta, Ontario, Quebec, and Newfoundland and Labrador. Since then, all provinces have implemented some form of indicators or standard assessment. These projects, along with SAIP, provide data for policy analysis with respect to some types of learning – primarily that associated with traditional curriculum goals (reading, writing, mathematics, science).

III. measuring change – Assessing changes in an individual's achievement is troubled by problems pertaining to the reliability and validity of tests. Considering that such changes are necessary, planned outcomes; inability to assess change is troubling. Recent developments in item response theory (May & Nicewander, 1998) and hierarchical linear modelling (Bryk & Raudenbush, 1992) offer viable solutions for tracking a student's progress.

IV. exploring levels of impact – Hierarchical linear modelling has been used in many contexts to ferret out achievement effects attributable to various levels of educational organization (e.g., classroom, school, district, province); use of such modelling in large scale assessments may increase the usefulness of the results when effecting change.

V. standards and expectations – Yackulic (1998) identified several problems that may limit the validity and generalizability of standards set by panels; Hambleton (1998) suggested strategies for improving standard setting procedures.

VI. consequential validity – large-scale assessments are intended to meet needs for accountability and system improvement. Messick's (1989) expanded notion of validity provides a framework for determining the impacts that actually occur.

Opportunity to learn

The provincial/territorial indicator projects and large-scale testing programs provide information on student performance that is useful to policy makers in describing differences among jurisdictions or between groups (e.g., males and females). However, to be used for improving educational outcomes, there needs to be additional information that can help explain variations in student performance. The concept of *opportunity to learn* (OTL) has been introduced to help explore some of the factors in the learning situation that may influence student achievement.

It is generally agreed that OTL can be organized into three broad categories:

1. *Personological factors* are those such as student ability, motivation, persistence, experience, or parent support.
2. *Classroom variables* include learning time, level of curriculum implementation, learning environments, teacher expertise, assessment practices or access to materials.
3. *Policy and resources allocation factors* are related to issues such as funding or equity policies.

Some of the earliest understanding of OTL was focused on *time spent on learning* as the primary OTL factor (Bloom, Carroll cited in Wang, Haertel &Walberg; 1990). Achievement was dependent upon the *amount of time* students spent in learning and on the *nature of the learning time* (i.e., engaged learning time). Porter (1993) examined the concept of OTL from the perspective of school delivery standards: i) the enacted curriculum "the actual content experienced in school" (p. 26); ii) effective pedagogy; and iii) actual resource materials for students. Later Porter (1995) returned to a discussion of the uses of OTL standards for either accountability and for school improvement; he envisioned support for school improvement as the more promising use of OTL.

Stevens (as cited in Wang 1998) defined OTL in the context of exposure to new knowledge. In this case, OTL is characterized by content coverage variables, context exposure variables, content emphasis variables and quality of a instructional delivery variables. Darling-Hammond and Falk (1997) addressed the question of OTL from the perspective of the learner. OTL is viewed as a specific set of delivery standards which include: adequate funding levels, access to well-prepared, fully-qualified teachers, reasonable class sizes, and access to the materials and equipment for learning (p. 196). Similarly, Guiton and Oakes (1995) emphasized the importance of OTL standards as a link between content (curriculum) standards and performance standards. It is their view that OTL is most useful when defined as access to new knowledge.

With respect to resource allocation, Howe (1994) expressed the view that it is doubtful whether additional resources alone can improve student performance. Hanushek (1997) also questioned whether there is a direct causal relationship between resource allocation and student performance. While (financial) resources are a necessary OTL factor, they are not sufficient to guarantee improved student performance (p. 153). On the other hand, Darling-Hammond and Falk (1997) and others suggested that OTL is largely a function of resource allocation by policy-makers.

One of the more recent approaches to OTL is to consider how OTL can be used to call attention to inequalities in student performance. Banks (1997) suggested that the concept of OTL, rather then being defined as a set of standards, could be used to begin examination and, by extension, explanation of the factors which account for inequalities in student achievement. One of the primary benefits to thinking about OTL is that attention is called to the differences in student

achievement. For example, OTL may help focus attention on inequalities of students' experiences in school or on the differences in the quality of instruction which students receive. Studies by Yoon and Resnick (1998) and Wang (1998) show some of the relationships between OTL and student achievement.

Yoon and Resnick examined the influence of student demographic information and teachers' classroom activities, teaching strategies and professional development on students' mathematics performance. An evaluation of the California Mathematics Renaissance program found evidence that teachers' instructional practices were positive influences on student achievement. The authors however, caution against over-interpreting such results and attributing causal effects.

In a similar study, Yang (1998) examined the effects of content exposure (as a measure of OTL) in grade 8 student science achievement. The study examined the effects of content coverage, exposure, and emphasis as well as the effects of five dimensions of instructional quality (teacher preparation, integration of concepts, material adequacy, equipment use and availability of textbooks). Student-level variables (attendance, gender, race, and pre-test scores) were controlled in the analysis. Results showed that attendance rate, content coverage and quality of instructional delivery were significant predictors of science achievement. The results also showed that OTL was a multi-dimensional concept. For example, teachers had greater impact on written tests than on hands-on tests.

Although OTL may provide an interesting perspective on explaining differences in group performance, caution is advised when using OTL in ways that are not defensible logically or psychometrically. For example, Stedman (1998) warned against simplistic OTL explanations of achievement variability on international assessments.

> "We should be wary, however, of explanations that attribute achievement differences solely to school factors or to a single organizational or instructional factor ... We should be wary of other claims that differences are primarily demographic and that school factors are largely irrelevant (p. 9)."

Although most investigations of OTL have been limited to more formal research studies, the concept is now being considered in applied research and testing projects. For example, Saskatchewan's recent Mathematics Assessment (Saskatchewan Education, 1998) assessments collected OTL information concurrent with measures of student performance. The OTL factors were measured using student self-reports. Results showed, for example, that students have insufficient opportunities for 'hands-on' math manipulatives in classrooms, that students exceeded expectations in terms of persistence in completing difficult mathematics problems, and that parent involvement in helping their children with school assignments exceeded expectations.

One of the limitations to research on OTL is that, as in the case of large-scale testing, the focus has been on a relatively narrow scope of learning outcomes (e.g., science or math). An expanded curriculum with an emphasis on affective and connative skills and talents will demand an expanded view of OTL. For example, what type of school experiences will provide opportunities to develop critical and creative thinking or social intelligence?

Conclusion

The provincial/territorial goals of education are a foundation for expanding the scope of classroom learning. Existing indicator and large-scale assessment programs have provided evidence that it is possible to develop and implement techniques to monitor student performance. However such monitoring is limited to a relatively narrow scope of intended learning outcomes: cognitive learning in traditional curriculum content. Current curricula are based on theories of cognition which emphasize not only knowledge acquisition, but procedural knowledge and higher-order thinking skills. Additionally, the curricula include explicit expectations with respect to the development of personal and social talents. Further, *opportunity to learn* is a factor that will help to better understand differences in student performance and to provide direction in making the policy changes to improve learning outcomes for all students.

The research questions below recognize that meaningful assessment of student learning must find a balance between classroom assessment and large scale testing.

A research agenda for the full scope of learning

1. In what ways can existing assessment practices be adapted to better address the full scope of learning?

   Presently large-scale tests focus on a relatively narrow conception of achievement. There is a need for research that adapts existing technology to measure goals related to the full scope of learning. Some specific questions could be:

   a. How can personal and social talents be assessed?
   b. Is it possible to establish benchmarks for such skills?
   c. What is the effect on instruction (and the more traditional learning outcomes) when non-traditional learning outcomes are the focus of large-scale assessment?
   d. How does one ensure the validity of measures of non-traditional learning?

2. In what way can emerging assessment technologies (e.g., computerized simulations) enhance the assessment of non-traditional outcomes? Can these technologies meet the need for realistic as situations argued for by situated cognition theorists?

3. To what extent does OTL assist in defining the full scope of learning and how can information on OTL assist current assessment practices?

4. In what ways are *assessment for accountability* and *assessment to improve student learning* linked?

   A common premise is that assessment for accountability can serve to improve education; research is needed to examine the nature of the link between assessment for accountability and assessment for improvement.

5. To what extent are existing pan-Canadian policies and practices congruent with curriculum goals and best-practice guidelines?

   All the provincial and territorial jurisdictions have endorsed the *Principles for Fair Student Assessment for Education in Canada.* With the rapid development of large- scale assessment programs, it is important for educators to know the extent to which such programs are congruent with those principles.

References

Anderson, J.R. (1990). *Cognitive psychology and its implications*. New York: W.H. Freeman.

Anderson, J.R., Reder, L.M. & Simon, H.A. (1997). Situative versus cognitive perspectives: Form versus substance. *Educational Researcher.* 26(1). 18-21.

Atlantic Provinces Education Foundation (1996). *Education indicators for Atlantic Canada.* Human Resources Development Canada

Banks, C.A.M. (1997). The challenges of national standards in a multicultural society. *Educational Horizons*. 75(3) 121-125.

Bateson, D. (1994). Psychometric and Philosophic Problems in "Authentic Assessment": Performance Tasks and Portfolios. *The Alberta Journal of Educational Research.* XL (2). 233-245.

Bereiter, C. (1997). Situated cognition and how to overcome it. In *Situated cognition: Social, semiotic, and psychological perspectives.* Kirshner, D.I, &. Whitson, J.A. (Eds). Mahwah, NJ, USA: Lawrence Erlbaum Associates, Inc. 281-300.

Bisanz, G.L., & Bisanz, J. (1994). Studying the development of academic skills in the 1990s: implications for assessment. *The Alberta Journal of Educational Research.* XL (2). 127-146.

Blackmore, D. (1998, October). Future of computerized testing. Paper presented at *Measurement and evaluation: current and future research directions for the new millenium.* Banff, AB.

Brown, J.S. (1988). *Situated cognition and the culture of learning*. Technical Report. BBN Labs, Inc., Cambridge, MA.; Xerox Corp., Palo Alto, CA. Palo Alto Research Center.

Bruner, J.S. (1966). *Toward a theory of instruction*. Cambridge, Ma: Harvard University Press.

Bruner, J.S., Goodnow, J.J., & Austin, G.A. (1956). *A study of thinking*. New York: Wiley.

Bryk. A.S., & Raudenbush, S.W. (1992). *Hierarchical linear models: applications and data analysis methods*. Newbury Park: Sage.

Canadian Education Statistics Council. (1996). *Education indicators for Canada*. Toronto.

Childress, H. (1998). Seventeen reasons why football is better than high school. *Phi Delta Kappan*. 79 (8). 616-619.

Cole, N.S. (1990). Conceptions of educational achievement. *Educational Researcher*. 19 (3). 2-6

Darling-Hammond, L. and Falk, B. (1997). Using standards and assessments to support student learning. *Phi Delta Kappan*. 79(3). 190-199.

Earl, L.M. (1995). Assessment and accountability in education in Ontario. *Canadian Journal of Education*. 20(1). 45-55.

Fagan, L.P. (1995). Performance accountability in the Newfoundland school system. *Canadian Journal of Education*. 20(1). 65-76.

Gagne, E.D., Yekovich, C.W., & Yekovich, F.R. (1993). *Cognitive psychology of school learning (2nd)*. New York: HarperCollins College Publishers.

Goleman, D. (1995). *Emotional intelligence*. New York: Bantam Books.

Greeno, J.G. (1989). A perspective on thinking. *American Psychologist*. 44(2). 134-141.

Guiton, G. and Oakes, J. (1995). Opportunity to learn and conceptions of educational equality. *Educational Evaluation and Policy Analysis*. 17(3). 323-336.

Hambleton, R. (1998, October). Advances in standard-setting methodolgy. Paper presented at *Measurement and evaluation: current and future research directions for the new millenium*. Banff, AB.

Hanushek A. (1997). Assessing the effects of school resources on students performance: an update. *Educational Evaluation and Policy Analysis*. 19(2). 141-164.

Hodgkinson, D. (1995). Accountability education in British Columbia *Canadian Journal of Education*. 20(1). 18-26.

Howe, K.R. (1994). Standards, assessment, and equality of educational opportunity. *Educational Researcher*. 23(8). 27-32.

Joint Declaration: Future Directions of the Council of Ministers of Education, Canada (CMEC) (1993). Victoria, BC.

Joyce, B. & Weil, M. (1972). *Models of Teaching*. Englewood Cliffs, NJ: Prentice-Hall.

Kelly, K.R. & Moon, S.M. (1998). Personal and social talents. *Phi Delta Kappan*. 79 (10). 743-746.

Lewin, K. (1935). *Dynamic theory of personality*. Trans. By D.K. Adams and K.E. Zener. New York: McGraw-Hill.

Maheu, R. (1995). Education indicators in Quebec. *Canadian Journal of Education*. 20 (1). 56-64.

Masters, G.N. and Misley, R.J. (1991). *New views of student learning*. Educational Testing Service. Princeton, NJ.

May, K. and Nicewander, W.A. (1998). Measuring change conventionally and adaptively. *Educational and Psychological Measurement*. 58 (6). 882-897.

Mayer, (1987). *Educational psychology: a cognitive approach*. Toronto: Little, Brown and Company.

McEwen, N. (1995). Accountability in education in Canada. *Canadian Journal of*

Education. 20(1). 1-17.

Messick, S. (1989). Validity. In R.L. Linn (Ed.), *Educational measurement* (3rd ed., pp. 13-103). New York: American Council on Education and Macmillan.

Misley, R.J.,Yamamoto, K., & Anacker, S. (1991). Toward a test theory for assessing student understanding. Educational Testing Service. Princeton, NJ.

Noonan, B.W. (1996). The five A's of school success. *The Canadian School Executive*. 16(1). 3-5.

Porter, A.C. (1993). School delivery standards. *Educational Researcher*. 22(5). 24-30.

Porter, A.C. (1995). The uses and misuses of opportunity to learn standards. *Educational Researcher*. 24(1). 21-27.

Principles for Student Assessment Practices for Education in Canada. (1993). Edmonton AB. Joint Advisory Committee.

Resnick, L. (1996). *Performance puzzles: issues in measuring capabilities and certifying accomplishments*. National Center for Research on Evaluation, Standards, and Student Testing. Los Angeles, CA.

Ryle, G. (1949). *The Concept of Mind*. London: Hutchinson.

Saskatchewan Education. (1998). *1997 Provincial Learning Assessment in Mathematics*. Regina.

Saskatchewan Education. (1997). *Saskatchewan Indicators Report*. Regina

Stedman, L.C. (1998). International achievement differences: an assessment of a new perspective. *Educational Researcher*. 26(3). 4-15.

Thorndike, E.L. (1912). *Education: A first book*. New York: Macmillan.

Tolman, E.C. (1932). *Purposive behaviours in animals and men*. New York: D. Appleton-Century.

Wang, J. (1998). Opportunity to learn: the impacts and policy implications. *Educational Evaluation and Policy Analysis*. 20(3). 137-156.

Wang, M.C., Haertel, G.D. & Walberg, H.J. (1990). What influences learning? A content analysis of review literature. *Journal of Educational Research*. 84 (1). 30-43.

Wales, C.E., & Stager, R. (1978). *The guided design approach*. Englewood Cliffs: Educational Technology Publications.

Yackulic, R.A. (1998, October). What is the nature of standard set by panels? Paper presented at *Measurement and evaluation: current and future research directions for the new millenium*. Banff, AB.

Yoon, B. and Resnick, L.B. (1998). *Instructional validity, opportunity to learn and equity: new standards for the California Mathematics Renaissance*. Center for Research on Evaluation,

Standards, and Student Testing. Los Angeles, CA.