Assessment - A Powerful Lever for Learning

Dr. Lorna Earl
Director, Aporia Consulting Ltd.

Abstract

Classroom assessment practices have been part of schooling for hundreds of years. There are, however, new findings about the nature of learning and about the roles that assessment can play in enhancing learning for all students. This essay provides a brief history of the changing role of assessment in schooling, describes three different purposes for assessment and foreshadows some implications that shifting to a more differentiated view of assessment can have for policy, practice and research.

Dr. Lorna Earl is Director, Aporia Consulting Ltd. and a recently retired Associate Professor in the Theory and Policy Studies Department and Head of the International Centre for Educational Change at OISE/UT. She is currently working on a contract as Researcher In Residence with the Ontario Ministry of Education. She has worked for over 20 years in schools and school boards and, as a leader in the field of assessment and evaluation, has been involved in consultation, research, evaluation and staff development with teachers' organizations, ministries of education, school boards and charitable foundations in Canada, England, Australia, Europe and the United States. Throughout her career, she has concentrated her efforts on issues related to evaluation of large-scale reform and assessment (large-scale and classroom) in many venues around the world.
A Brief History of Assessment

Formal and informal assessments of learning have existed for centuries from the early Chinese civil service exams for entry into high public office, public presentations by Aristotle’s students, and practical assessments for entrance to medieval guilds. But it wasn’t until the advent of industrialization and universal schooling at the turn of the last century that schools became significant social institutions and the evaluation of student achievement, as we know it, became a significant dimension of universal schooling.

Over the years, assessment in schools has largely been “summative”, designed not only to report the achievement of individuals to parents, other teachers and the students themselves, but to provide the evidence for decisions about students’ placements and life choices. Students have been given different programs, sorted into different “tracks” and set on the voyage towards their various destinations based on a system of assessment and reporting. Teachers have been the gatekeepers and tests and examinations have been the “gates” students have to pass through to move to the next level of education.

About the middle of the twentieth century, it became clear that schooling was an important key to social mobility and that achievement in school was the basis for entry into the workplace. Tests and exams took on major importance as the basis for deciding which students would have access to additional educational opportunities. Many jurisdictions instituted standardized testing programs alongside classroom assessment to ensure fair, accurate, and consistent opportunities for students. Assessment was considered a mechanism for providing an index of learning, and it followed a predictable pattern: teachers taught, the students’ knowledge of the material was tested, judgments about the students were made based on the testing, and students moved on to the next unit or grade or school.

During the 1970s and 1980s, the purposes for classroom assessment expanded. The terms formative assessment and summative assessment entered the language of educators, formative assessment being assessment that takes place during teaching to make adjustments to the teaching process and summative assessment being assessment at
the end of a unit or term to convey student progress. In order to fulfill these two different purposes, educators extended their assessment practices and began assessing a wider range of student work such as practical tasks, coursework, projects, and presentations. For the most part, however, assessment was still a matter of making statements about students’ weaknesses and strengths.

More recently, the focus in educational policy has been on preparing all students for tomorrow’s world and the expectations for students have increased in breadth and depth. Completion of secondary school is becoming a basic requirement for entering the workforce.

At this point in time, schools and districts are serving contradictory purposes of “education for all” and “education as gatekeeper”. Teachers and administrators are the instruments of these contradictory demands and are both recipients and perpetrators of these competing messages. In this confused and emotionally-charged assessment environment the stakes are high to “get it right”. Educators find themselves in a difficult position. They are part of the transition, laden down with the burdens of the past, while contemplating the possibilities of the future. They know how it has always been and have a great deal invested in maintaining stability but, at the same time, many of them acknowledge that it just doesn’t feel right.

Assessment is a pivotal concept in this more complex learning environment. It has been the primary mechanism for the gate keeping role of schools. We are also becoming aware of the tremendous power of assessment in learning.

Assessment And Learning

One of the most promising recent contributions from educational research has been the work done to understand the role of assessment in the process of learning. In a landmark paper, Black and Wiliam (1998) synthesized over 250 studies linking assessment and learning and found that the intentional use of assessment in the classroom to promote learning improved student achievement. This meta-analysis supported the comprehensive review done by Crooks (1988) in which he made a strong research-based case that classroom assessment had
both short- and long-term effects on learning. In the short term, classroom assessment could:

- focus attention on important aspects of the subject,
- give students opportunities to practice skills and consolidate learning, and
- guide further instructional or learning activities.

In the medium and long term, assessment held the possibility of:

- influencing students’ motivation as learners and their perceptions of their capabilities,
- communicating and reinforcing teaching goals, including performance criteria and desired standards of performance,
- influencing students’ choice of and development of learning strategies, skills and study patterns, and
- influencing students’ subsequent choice of courses, activities and careers.

It has become obvious that assessment is a powerful tool for learning.

If learning is the goal, dramatic changes are required in the way that assessment is used in classrooms. Learning was long thought to be an accumulation of atomized bits of knowledge that are sequenced, hierarchical, and need to be explicitly taught and reinforced. Learning is now viewed as a process of constructing understanding by attempting to connect new information to what is already known so that ideas have some personal coherence. Costa (1996) identified nine uniquely human strategies for harnessing information and making sense of the world through an intentional process of learning:

- **Metacognition**: Human beings can reflect on their own thinking processes. Experts describe their thinking as an internal conversation – monitoring their own understanding, predicting their performance, deciding what else they need to know, organizing and reorganizing ideas, checking for consistency between different pieces of information and drawing analogies that help them advance their understanding.

- **Constructing abstraction**: Humans have the capacity to use language, images and numbers as symbols to
transform events into categories and patterns. These symbolic systems make it possible for people to think in abstractions and to order and reorder the world in thought.

- **Storing information outside the body**: Humans store, organize and retrieve information in and from locations other than their bodies. From cave drawings, to books, to videotdisks, external storage and retrieval systems provide access to information far beyond the limits of memory.

- **Systems thinking**: Humans are able to see patterns, congruencies, and inconsistencies while still focusing on the whole. This capacity allows them to consider many perspectives and to imagine how changing one element can have an impact on the total system.

- **Problem finding**: Not only are humans able to search for problems to solve; they appear to enjoy it. Humans question and sense ambiguities and anomalies in the world around them. Once there is some doubt, they look for better ways of understanding the nature of things.

- **Reciprocal learning**: Human beings are social creatures with a compulsive craving to engage with each other. They learn best in groups, as they listen to one another, strive for agreement and rethink their beliefs and understanding.

- **Inventing**: Human beings are creative and often motivated intrinsically, rather than extrinsically, to work on tasks because of the challenge. They constantly strive for greater fluency, elaboration, novelty, parsimony, simplicity, craftsmanship, perfection, harmony, beauty, and balance.

- **Deriving meaning from experience**: One of the most significant attributes of human beings is that they can reflect on and learn from their experiences. They can stand back, monitor activities and modify actions or beliefs.

- **Altering response patterns**: Although a certain amount of human activity may be hard-wired, people are able to make significant conscious and deliberate choices about
their behaviour. They are always capable of learning and altering their responses based on new ideas or understanding.

Costa’s framework accentuates the view that learning is not a static trait; learning is a dynamic process that itself can be learned and developed. Humans have many different ways of approaching something new, of investigating it and of making sense of it. Viewed this way, learning is an ongoing, iterative process of fitting information into patterns or schema of similarities, differences, likeness and regularities.

Assessment in this conception of learning is much more than a summary or index of learning, at the end. It is an integral part of the learning process. When teachers use classroom assessment to become aware of the knowledge, skills, and beliefs that their students bring to a learning task, use this knowledge as a starting point for new instruction, and monitor students’ changing perceptions as instruction proceeds, classroom assessment promotes learning.

The Assessment Reform Group (1999) in England called this process Assessment For Learning and assessment specialists in many countries have been engaged in a number of studies and development projects designed to understand it better (Gipps & Cummings, 2003; Earl & Katz, 2005; Timperley & Parr, 2003; Little et al., 1991; Wiggins & McTighe, 1998; Stiggins, 2003).

Teachers who are assessing for learning use assessment to identify particular learning needs and draw on this information to ensure that pupils get on track in their learning, the assessment process happens throughout learning and teaching, not at the end.

**Purpose Is All**

Classroom assessment has always been used for a variety of purposes but these purposes are becoming more differentiated and complex. In recent publications, colleagues and I haveexpanded on the differentiation made by the Assessment Reform Group to describe three intertwined but distinct assessment purposes – *assessment for learning; assessment as learning; and assessment of learning* (Earl,
Assessment - A Powerful Lever for Learning

2003; Earl and Katz, 2005).

Assessment for learning is designed to give teachers information that will allow them to modify the teaching and learning activities in which students are engaged, in order to differentiate and understand how individual students approach their learning. Such assessment suggests that students are all learning in individual and idiosyncratic ways, while recognizing that there are predictable patterns and pathways that many students go through. The information from carefully-designed assessments (and critical reflection on the assessment process) can be used to determine not only what students know, but also to gain insights into how, when, and whether students use what they know, so that teachers can streamline and target instruction and resources.

Assessment as learning is a subset of assessment for learning that emphasizes using assessment as a process of developing and supporting metacognition for students. Assessment as learning focuses on the role of the student as the critical connector between assessment and learning. Students, as active critical thinkers make sense of information, relate it to prior knowledge, and use it for new learning. This is the regulatory process in metacognition. It occurs when students personally monitor what they are learning and use the feedback from this monitoring to make adjustments, adaptations and even major changes in what they understand. When teachers focus on assessment as learning, they use classroom assessment as the vehicle for helping students develop, practice and become critical thinkers who are comfortable with reflection and the critical analysis of their own learning.

Assessment of learning is assessment used to confirm what students know, to demonstrate whether or not the students have met the standards and/or show how they are placed in relation to others. In assessment of learning, teachers should concentrate on ensuring that they have used assessment to provide accurate and sound statements of proficiency or competence for students, so that the recipients of the information can use the information to make reasonable and defensible decisions.
Assessment For Learning

When learning is the goal, teachers and students collaborate and use ongoing assessment and pertinent feedback to move learning forward. When classroom assessment is frequent and varied, teachers can learn a great deal about their students. They can gain an understanding of students’ existing beliefs and knowledge and can identify incomplete understandings, false beliefs, and naïve interpretations of concepts that may influence or distort learning. Teachers can observe and probe students’ thinking over time and identify links between prior knowledge and new learning. Teachers are aware that their classes are made up of students with different needs, backgrounds, and skills. Each student’s learning is unique and the context of classrooms, schools, and communities can be very different for each student. As well, the societal pressure for more complex learning for all students necessitates that teachers find ways to create a wide range of learning options and paths so that all students have the opportunity to learn as much as they can, as deeply as they can, and as efficiently as they can.

In the past, instruction and assessment were differentiated only for those students with identified needs. The class was typically regarded as a homogeneous unit and teachers used phrases such as “The lesson went well for the class” or “My students seemed to grasp that concept well.” Any student for whom the lesson did not go well was considered an exception. Within this framework, teachers accounted for the variability in their students’ performances on the basis of particular dimensions like “Learning Disability,” “English as a Second Language,” “Attention Deficit Hyperactivity Disorder,” or “gifted.” Students with these labels were seen as “different” from the rest of the class, and the rest of the class was seen as a single entity.

Although such labels are almost irresistible, they are not an explanation for why some students are not learning. Differences exist among all students. Individuals, not classes, learn. Classes, as teachers know, are made up of individual students with their varied interests, backgrounds, and understandings. Assessment is the tool that teachers can use to understand these differences and tailor their instruction to particular students and their learning patterns.
Assessment As Learning

Learning is also enhanced when students are encouraged to think about their own learning, to review their experiences of learning (What made sense and what didn’t? How does this fit with what I already know, or think I know?), and to apply what they have learned to their future learning. Assessment provides the feedback loop for this process. When students (and teachers) become comfortable with a continuous cycle of feedback and adjustment, learning becomes more efficient and students begin to internalize the process of standing outside their own learning and considering it against a range of criteria, not just the teacher’s judgment about quality or accuracy. When students engage in this ongoing metacognitive experience, they are able to monitor their learning along the way, make corrections, and develop a habit of mind for continually reviewing and challenging what they know.

Teachers can teach students to be thoughtful, systematic and rigorous in reflecting on their own learning and making adjustments so that their learning is more efficient and makes sense for them.

Assessment Of Learning

Assessment of Learning is also an important assessment purpose because it is the place where students’ learning is public and accessible to others. It is also very “high stakes” for students and needs to be accurate and fair, based on evidence captured from a variety of contexts and applications. Effective assessment of learning requires:

- a rationale for why it is necessary to undertake a particular assessment of learning at any point in time
- clear descriptions of the intended learning
- processes that make it possible for students to demonstrate their competence and skill
- a range of alternative mechanisms for assessing the same outcomes
- public and defensible reference points for making judgments
- transparent approaches to interpretation
Changing Assessment: Implications for Policy, Practice and Research

Assessment is a highly emotional and political activity that sits at the heart of the work of schools. Just identifying and naming the different purposes for assessment will not reduce the conflicts among them or create a context within which the differences will be understood and respected. Instead, the road to assessment reform that unleashes the power of assessment as a lever for learning will be a winding one, with implications for policy, for practice, and for research.

Implications for Policy

Assessment is part of a much bigger set of challenges within schooling. As Broadfoot (2001, 109) identified: “Current developments in society are set within a context that is characterised by a confused muddle of institutions and practices that are the enduring product of previous eras.” Assessment, as she points out, is one of the muddled processes in transition from the past.

Government policies routinely place conflicting demands on schools. Much of the debate (both overt and hidden) focuses on the purpose of schooling. Some legislation assumes that schools are instruments of social control that need to be controlled themselves. Other legislative directions identify schools (with rhetoric about fairness and equity) as a mechanism for maximizing opportunities for all. The increasing focus on accountability is thus combined with a societal push towards enhanced learning for all. Policy makers often avoid choosing between these different value positions about the purpose of education by blurring the issues and appealing to both camps.

Policy makers use and recommend assessment for multiple purposes – supporting learning; reporting achievement of individuals
Assessment - A Powerful Lever for Learning

for certification, progress and transfer; and satisfying the demands for public accountability. These purposes are often in conflict with one another, although both are legitimate. The tensions between these purposes have important implications for choices about the agencies which should conduct assessments and the optimum instruments and appropriate interpretations to serve each purpose. Different purposes require different approaches and the attempt to mix the purposes is likely to ensure that none of them will be well served. It is important to decide the purpose first and then design an assessment program to fit (Gipps, 1994).

Differentiation in relation to purposes for assessment means that policy-makers must realize and acknowledge that they are making different and competing demands for the use of assessment in education. This recognition would necessitate clarity about which purposes are being addressed within various policy environments or decisions and ensuring that the assessment approaches are consistent with the purposes.

If they are going to make assessment “fit for purpose”, policy makers cannot sit on the fence about why and how they use assessment in decision-making and that they cannot rely on assessments to fulfill these competing purposes at the same time. The purpose of assessment determines the nature of the assessment process and the way the results are presented, interpreted and used. Policy makers have a responsibility to lead the way not only in “naming” the different purposes of assessment, but also in differentiating assessment approaches so that the results are appropriate and adequate to the particular purpose at hand and are not used inappropriately for purposes that they are not capable of supporting.

Implications for Practice

Changing assessment practices in schools requires some deep conceptual changes on the part of educators and fundamental shifts in thinking about teaching and about assessment and about the relationship between them.

As concepts, assessment for learning and assessment as learning are rational and persuasive ideas that might appear to require
straightforward changes – adjusting a small part of what teachers do (their classroom assessment practices) to focus more explicitly on pupil learning. In fact, rethinking assessment challenges many of the taken-for-granted precepts of schooling and threatens to reconfigure both the purpose and the culture of schools. These shifts in assessment are rooted in a belief that all students can learn, and that the teacher’s role is to find a range of ways to ensure that this learning happens. Assessment becomes an integral part of the learning-teaching process, as the feedback loop for learning, with the emphasis in many assessment events shifting from making judgments that categorize students, to using them as windows into learning. This is a new way of thinking about assessment for most teachers that runs counter to the long-standing histories and experiences of most teachers.

Because assessment is very public for pupils, parents and teachers, shifting assessment also has serious implications for interaction and communication between teachers and parents, as well as students. Educators can not change assessment in dramatic ways without including parents and students in their thinking and helping them understand and engage with the purposes and the approaches that they are going to be part of.

Even though teachers are increasingly aware that assessment has multiple purposes and are trying to include assessment for learning and assessment as learning in their daily work, they find the shift uncomfortable and difficult. They are caught between the competing purposes of classroom assessment and are often confused and frustrated by the difficulties that they experience as they try to reconcile the demands. Redefining assessment will require concerted effort on the part of teachers (Katz, 2002) and considerable support from administrators and the public more broadly.

**Implications for Research**

As the power of assessment for learning becomes clearer, researchers have begun focusing on the application of this knowledge into actual classroom practice. Realizing the potential of assessment will depend on how the public, policy makers and educators come to understand and to use assessment. There is much to learn about how
education can harness and use assessment most effectively to ensure high level learning for all.

There are innumerable questions about the nature of the policy environment and the public messages that surround assessment both large-scale and classroom. For example, what purposes are actually embedded in policy text and how is it communicated and understood by the recipients?

There are also questions for investigation at the level of school administration. What role do school leaders play in capacity building for rethinking assessment with purpose in mind? What kinds of reporting to parents provide the most information to support learning for students?

Within classrooms, there is a research program developing about how different assessment practices work to enhance learning. Clarke (2001) has been engaged in a program of research directed at understanding feedback and how it works in literacy learning in primary classrooms. In her research she has found that feedback is most useful when it focuses on the learning intentions of the task and when students get differentiated feedback, depending on the learning that they bring to the particular task. Black et al., (2003) have been investigating the role of questioning as a form of classroom assessment and how it can reinforce learning. Many questions remain to be addressed. For example, how does assessment contribute to grouping and regrouping in classrooms for differentiated learning? What is the role of peer assessment in enhancing learning?

Summary

Classroom assessment is a complex and intricate process that has enormous power to transform learning for all students. But, assessment has a long history and the transition to more differentiated understanding and application of classroom assessment practices will be a challenging one. Policy makers, practitioners and researchers need to stand back and view assessment through new lenses, be prepared to let go of some of their deeply held notions about what assessment is for, use what is known to reconfigure the assessment process, and engage in
ongoing inquiry to learn more.

References


Assessment - A Powerful Lever for Learning


The Conundrum of Classroom Writing Assessment

Drs. Joseph F. Engemann and Tiffany Gallagher
Brock University

Abstract

This article highlights the need for reliable, valid, and fair assessment of writing performance by classroom teachers to balance large-scale literacy assessment initiatives that have taken root in most jurisdictions across Canada. While these initiatives have established a focus on writing instruction and assessment, there is a notion that large-scale assessment may not, by itself, provide an accurate reflection of students’ writing performance. It is important, therefore, for teachers to use effective assessment instruments, such as rubrics, to reliably, validly, and fairly determine how well their students write. The grading of written compositions through the use of rubrics can be effectively implemented by teachers with sufficient training whilst providing detailed student feedback. Classroom assessment of writing must therefore be coupled with professional development and teacher collaboration, which should lead to an improvement in student writing proficiency.

Dr. Joseph Engemann is a member of the Department of Graduate and Undergraduate Studies in Education at Brock University. He specializes in research on classroom assessment and evaluation, teacher induction and mentoring, and aboriginal science education.

Dr. Tiffany Gallagher is a member of the Department of Pre-service Education at Brock University. Her current research interests include literacy assessment, reading and writing strategy instruction, teachers with learning disabilities, and the role of the in-school resource teacher.
Long regarded as one of the “three Rs,” along with reading and arithmetic, writing is considered to be one of the essential academic skills that students acquire as an intended function of schooling (Barakett & Cleghorn, 2000). In keeping with this commonly accepted notion, there is a current culture of accountability in education, which emphasizes standards and evaluation of these basic academic skills (Hunter, Jones, & Randhawa, 1996). In all Canadian provinces and territories except Prince Edward Island and Nunavut, large-scale writing assessment is conducted either annually or over a two- or three-year period as a way of measuring the writing proficiency of students (Airasian, Engemann, & Gallagher, 2007). While claiming to address the need for public accountability, it has been argued that large-scale assessment is woefully inadequate at providing the detailed diagnostic evidence and formative recommendations needed to improve classroom pedagogy and student achievement (Froese-Germain, 1999; Simner, 2000).

According to the Standards for the Assessment of Reading and Writing (IRA/NCTE Joint Task Force on Assessment, 1994), valid assessment of writing should involve multiple perspectives and sources of data. Generally, large-scale writing assessment relies on a ‘one-off’ writing sample that is used as the reflection of students’ writing proficiency (IRA/NCTE Joint Task Force on Assessment, 1994). This single writing opportunity places pressure on classroom teachers to coach their students for these assessments (Mabry, 1999). Some teachers adjust their instructional practices and calibrate their writing assessment tools to approximate the protocols of large-scale assessments (Hunter et al., 1996). Other teachers find themselves grading students’ writing samples and providing them with constructive feedback specific to the criteria of the large-scale assessments (Earl, 1999; Skwarchuk, 2004).

Undoubtedly, teachers are eager to undertake practices that will provide them with the most accurate and efficient way of assessing their students’ writing. The focus of this article, therefore, is to provide an overview of classroom writing assessment with a focus on the use of rubrics. The need for teachers to ensure that writing assessment is reliable, valid, and fair will be emphasized.
An Overview of Classroom Writing Assessment

Writing may be considered one of the most difficult proficiencies that students are expected to master in elementary school (Gunning, 2002). Furthermore, classroom teachers typically find assessment of students’ writing performance onerous, time-consuming, and subject to the vagaries of imprecision and subjectivity (Culham, 2003). As a result, there is a resurgence of interest in models of instruction that assist teachers in defining the components of good writing and that provide a structure for assessment that connects with identified characteristics (Culham, 2003; Spandel & Hicks, 2002). Currently, many teachers are searching for more reliable and valid methods of assessment of writing performance that they can employ as part of their regular classroom practice.

Over the past four decades, teachers have tended to assess writing with holistic grading methods (Hunter et al., 1996). Holistic grading involves making an assessment of the quality of a complete written composition against a prepared scale or rubric (Hunter et al., 1996). It involves assigning a global grade (percentage, rating, letter, etc.) as a measure of the students’ level of writing performance. Holistic grading for writing is appropriate when the purpose of assessment is to obtain a broad perspective about the writing proficiency of a student or when the written composition cannot be assessed according to distinct criteria (Moskal, 2000). It is preferred by many teachers as a quick and efficient way of assessing students’ writing. By contrast, analytic grading involves the breaking down of a written composition into components, each of which is assessed separately and then amalgamated with the scores from other components to derive an overall grade. Analytic grading can provide a more comprehensive outline of the strengths and weaknesses of students’ writing performance than holistic grading, but it is detail-oriented and thus more time-consuming.

Reliable, Valid, and Fair Writing Assessment

Concerns about the reliability and validity of writing assessment are ever-present in the minds of teachers. For example,
some educators regard holistic grading as lacking uniform precision since there is a requirement to globally judge students’ writing compositions. Indeed, without precise assessment tools, teachers may assess written compositions subjectively and inconsistently. Yet, there is evidence that holistic scoring measures can be highly reliable when graders are extensively trained in the application of the measures. In fact, in these cases, there can be strong correlations between holistic and total analytic scores (Hunter et al., 1996). Regardless of the type of method used to grade writing, if there is a lack of alignment between writing instruction and assessment, teachers may be led to assess components that were not included as an instructional focus. Teachers need to look long and hard at assessment measures to ensure that the outcomes described reasonably reflect their instructional objectives. Effective writing assessment protocols must be consistent, accurate, and reasonable or, in other words, reliable, valid, and fair.

**Reliability**

Reliability refers to the degree of stability or consistency of assessment among multiple raters at one point in time or a single rater at different points in time (Airasian et al., 2007). It also refers to the accuracy of an assessment tool itself (Isaac & Michael, 1997). The goal in assessment is to establish a consensus in the application of an assessment tool. In writing assessment, this may be accomplished through the use of a tool that has clearly articulated criteria and the implementation of in-service training sessions to establish high levels of inter-rater reliability (Stuhlmann, Daniel, Dellinger, Denny, & Powers, 1999). Ideally, these sessions should encourage teachers to hone their skills in consistently applying assessment tools.

**Validity**

Validity is the extent to which assessment data are appropriate for making a decision about students and instruction (Airasian et al., 2007). An assessment tool must have content validity; that is, it must contain criteria and descriptors that accurately describe the students’ products or performances. In classroom settings, an important
consideration is instructional validity, which is the degree to which the assessment measures the classroom instruction (Santrock, Woloshyn, Gallagher, DiPetta, & Marini, 2004). This implies that assessment tools should measure what was taught and whether students have had adequate opportunity to learn the skills and obtain the knowledge that is being assessed. At times, grading students’ written compositions requires evaluative judgment and there is a temptation to subjectively utilize assessment tools (Aiken, 2000). To reduce teachers’ misuse and misinterpretation of writing assessment measures, they must be provided with assessment standards that are easily identifiable and can be applied to every student in accordance with instructional objectives (Banks, 2005).

**Fairness**

In general, fair and equitable assessment is the ethical responsibility of all teachers. In Canada, the criteria for determining fairness and equity in student assessment have been outlined as principles and guidelines in *Principles for Fair Student Assessment Practices for Education in Canada* (Joint Advisory Committee, 1993). This document was produced through a collaborative effort among many recognized organizations within the Canadian educational community and has been supported by the Canadian Teachers Federation, the Canadian School Boards Association, and the Canadian Society for the Study of Education. This framework outlines principles for classroom assessment that guide teachers in developing and choosing methods for assessment, collecting assessment information, judging and scoring student performance, and summarizing, interpreting, and reporting assessment findings (Joint Advisory Committee, 1993). An extrapolation of this framework into the assessment of written compositions means that teachers must adopt reasonable standards of writing assessment while, at the same time, maintain respect for students’ individuality in writing characteristics such as tone or voice. In addition, the IRA/NCTE Joint Task Force on Assessment (1994) states that equity should be pursued in writing assessment, a goal that is unlikely to be achieved through the use of a single assessment. Instead, several varied assessment pieces are
required to provide a meaningful measure of the writing capabilities of students.

The Birth of the Rubric

One of the purposes of assessment is to encourage students’ skill development by providing evaluative feedback; in writing assessment, rubrics are particularly effective for this purpose (Linn & Miller, 2005). A rubric is a set of criteria for different levels of performance (Airasian, 2005; McMillan, 2004). Rubrics, like the grading strategies that use them, may be identified as either holistic or analytic. A holistic rubric focuses on grading in a more general and overall fashion as the written composition is assessed in its entirety (Linn & Miller, 2005). By contrast, analytic rubrics focus on particular characteristics of the writing and pinpoint strengths and weaknesses of a composition based on specifically delineated criteria (Linn & Miller, 2005). Within an analytic rubric, the criteria represent the characteristics that help the rater identify the goals of instruction, while the descriptors identify the degree to which the student has attained these goals (McMillan, 2004). The descriptors provide illustrative detail about the criteria. The number of characteristics within a rubric can vary and the objective description of these criteria is often difficult for the rubric author to document (Linn & Miller, 2005).

To choose or construct an appropriate rubric, educators should consider the assessment purpose (Linn & Miller, 2005), measurement technique(s), and the implications of the measurement tool on evaluation (McMillan, 2004). High-quality rubrics have at least four characteristics: (1) content/coverage of the features that are used to measure the quality of the performance or product; (2) clarity/detail of the definitions, indicators, and samples of student work; (3) practicality of the rubric in terms of its usefulness for instruction and assessment; and (4) quality/fairness, which considers the degree of inter-rater reliability and fairness to all students (Arter & McTighe, 2001). In addition to these characteristics, a rubric should clearly connect instruction with assessment through its evaluative criteria (Payne, 2003).

As noted above, rubrics may be used for scoring or instruction
The Conundrum of Classroom Writing Assessment

(Andrade, 2005). Scoring rubrics are used strictly to assign grades, whereas, instructional rubrics are used during the instructional process. Teachers use instructional rubrics to clarify learning goals and focus their teaching. In this fashion, an instructional rubric is provided to students along with an assignment and then used for formative feedback and summative evaluation. A teacher’s formative feedback is often specific to rubric criteria and allows students to target aspects of their performance that require improvement. In essence, by providing students with assessment criteria, teachers make learning targets clear to students, which should improve their writing performance (Arter & McTighe, 2001).

**Contemporary Rubric Use**

Over the past decade, analytic trait-based rubrics have grown in popularity among educators in North America (e.g., 6 + 1 TRAIT model, Culham, 2003; Write Traits®, Spandel & Hicks, 2002). These rubrics provide classroom teachers with specific measures to assess the six componential writing traits: ideas, organization, voice, word choice, sentence fluency, and conventions. Each trait is individually assessed on a 5-point scale allowing teachers the facility to provide students with specific feedback on separate traits in terms of strengths and weaknesses. In general, this kind of formative feedback is especially effective at improving students’ learning (Black & Wiliam, 1998). Most particularly, analytic trait-based rubrics assist struggling writers to examine all aspects of good writing, not just those related to their specific need for growth and development (Schirmer & Bailey, 2000). More research needs to be conducted, however, to determine the veracity of the claims about the effectiveness and the reliability of analytic trait-based rubrics that are used to assess writing performance.

**Reliability of Writing Rubrics**

Contemporary forms of writing assessment raise the concern about whether different graders can reliably assess writing samples using complex scoring criteria. Consequently, an integral aspect of establishing reliability is to ensure that graders understand the criteria
that they are using to assess student performances and products, and
that they come to a consensus as to how the criteria are to be applied.
This process is one of negotiation and involves determining inter-rater
reliability among graders. Inter-rater reliability is a measure of the
degree of correlation of responses among graders who are assessing a
single product or performance. It is an essential characteristic of any
assessment that will be used for placement or promotion and for the
differentiation of instruction among learners (Aiken, 2000). For writing
assessment, an inter-rater reliability score is represented by an
intraclass correlation coefficient, a generalized inter-rater measure
derived from several graders scoring the compositions of a number of
students (McGraw & Wong, 1996). The process of establishing inter-
rater reliability among teachers requires an investment of time through
professional development.

As an illustration, 10 literacy resource teachers participated in
two consecutive 3-hour sessions aimed at discussing elementary
writing assessment and evaluation (Gallagher, Mangat, Engemann, &
Castle, 2005). Three scoring rubrics (Write Traits®, Spandel & Hicks,
2002; 6 + 1 TRAIT model, Culham, 2003; Wechsler Individual
Achievement Test [WIAT], The Psychological Corporation, 1991) were
used to establish an intraclass correlation coefficient for the grading of
writing samples. A comparison was made among the teachers’ total
grades for each of the papers. Intraclass correlation coefficients were
calculated for the Write Traits® rubric, $r = 0.94, p<.05$; for the 6 + 1
TRAIT rubric, $r = 0.94, p<.05$; and for the WIAT rubric, $r = 0.96,
p<.05$. In the end, the teachers were able to reliably use each of the
three rubrics to score a series of writing samples. However, it should
be noted that this result came after a substantial amount of discussion
and negotiation among the raters about the meaning of the descriptors
and qualifiers in the rubrics.

As they graded the student writing samples, the resource
teachers openly provided comments about the efficacy of the rubrics
(Gallagher et al., 2005). Initially, the resource teachers found it
difficult to interpret the exact meaning of some of the adjectives used in
the rubric descriptors. Through the process of discussing concerns,
contrasting rubrics, debating grading decisions, and supporting
judgments, the resource teachers struck a consensus on the application
of the scoring rubrics. Across the two training sessions, cohesion developed among the resource teachers and they perceived that their discussion contributed to the establishment of consistency in their grading. Consensus-building around the use of rubrics would appear to lessen the subjectivity that permeates the assessment of writing. This happens as raters negotiate the meaning of the descriptors within a rubric.

An important mechanism for developing consensus among teachers is through collaborative professional development. Engemann, Gallagher, and Castle (2005) found that most teachers eagerly seek to improve their instructional practice and assessment of student work either through attendance at workshops or by informal collaboration, mentoring, and dialogue with colleagues. Effective professional development has the benefit of helping teachers gain consistency in their assessment of student writing through the development of similar pedagogical practice. Rubrics are more reliably employed when teachers are trained to apply them consistently during assessment. Dialogue and debate are essential to broadening teachers’ understandings of the elements of writing and the components of writing assessment.

**Future Classroom Writing Assessment**

Teachers can improve the reliability in their use of rubrics for grading written compositions by enlisting the support of colleagues. For example, teachers may ask other teachers to grade one or more compositions from students in their classrooms and compare the grading. Alternatively, teachers can receive collaborative professional development in the application of grading criteria. This should lead to higher rates of agreement among different graders with the accompanying likelihood that the rubric will be more reliably applied in assessment (IRA/NCTE Joint Task Force on Assessment, 1994). This is especially important for the use of analytic rubrics. As an illustration, primary-level teachers who were trained to interpret the scoring criteria on an analytic writing rubric, graded written compositions more consistently than untrained teachers (Stuhlmann et al., 1999). The teachers who received training in the use of a writing
Engemann and Gallagher

assessment rubric tended to be more uniform in their interpretation of criteria within the rubric. However, teachers should not lose sight of the fact that reliable assessment is not necessarily valid assessment. The assessment task must be meaningful to draw worthy conclusions. Furthermore, writing assessment should be ongoing and reflective of the classroom writing instruction over a period of time.

Conclusion

The need for an improvement in student writing proficiency is evident to some who teach in institutions of post-secondary education. Many students are graduating from high school with a less than adequate level of writing proficiency. Teachers, teacher educators, and literacy consultants must assist each other in providing opportunities for all educators to improve the quality of writing instruction and assessment through collaboration and meaningful professional development.

References


Scarborough, Canada: Prentice-Hall Canada, Inc.


Pedagogical Synergy: Linking Assessment, Curriculum, and Instruction

Drs. Blair Mascall and Carol Rolheiser
Ontario Institute for Studies in Education,
University of Toronto

Abstract

This paper describes the evolution of attempts to build coherence and capacity in an Ontario school district, focusing on the development of literacy strategies in all of the district’s elementary and secondary schools. In reviewing case studies in four elementary schools, the authors have identified three key elements (instruction, curriculum, and assessment) as the key dimensions which have the greatest influence on student achievement. The authors of this paper present a new construct, pedagogical synergy, in which those three elements are combined. Improvements can occur at both the district and school levels when there are horizontal and reciprocal strategies for building capacity and increasing coherence. It is the mutual support between district and schools that provides the power in this new concept.

Introduction

When we started our research in the York Region District School Board (YRDSB) in 1999, we were exploring the development of capacity in the implementation of an early literacy pilot program. We use the term capacity in this paper to describe the knowledge and skills of individuals or groups. We used a conceptual framework by Newmann, King and Youngs (2000) which described the relationship between government or district policies and programs and student
achievement, by identifying two key variables: instructional quality, (which is the major determinant of student achievement), and school capacity, which the authors postulated was the determinant of instructional quality (see Figure 1). We found this a helpful model as we focused on important dimensions of school capacity. Our subsequent research has moved us beyond that model, to conceptualizing a model of district capacity. We often return to the Newmann, King and Youngs research, however, as a foundational piece in our understanding of collective capacity at the school level.

![Diagram of Policies & Programs]

**Figure 1.** Policies and Programs

In our latest research in the York Region District School Board (see Mascall, Rolheiser, Wallace, Anderson, & Fullan, 2005), we found that three elements continued to come into the foreground. These were the elements of Instructional Quality, which Newmann, King and Youngs (2000) identified as curriculum, instruction and assessment. In
this paper, we examine the relationship between these three elements, with a particular focus on the role of assessment in strengthening school and district capacity. This analysis leads us to the development of a new model, focusing on the relationship amongst these three key elements, in what we term ‘pedagogical synergy.’

The York Region District School Board

YRDSB serves over 115,000 students north of Toronto in 167 schools: 27 of which are secondary and 140 elementary. The number of elementary students has grown 4.8% per year in the five years since 2001. Forty new schools have been opened since 2000, and the school District has announced that it will open more new schools in the next few years to meet the demands of a growing population. A number of existing elementary schools are filled to capacity. This rapidly growing District serves a diverse socio-cultural and linguistic population, with over 100 languages spoken in the schools. There are 8,811 teachers in the system. Many district-wide initiatives are in place in the YRDSB, three of which are described here. The first of these programs, the Assessment Training Program (ATP), is a comprehensive initiative established in 1998. The goal is to build understandings about effective diagnostic, formative, and summative assessment practices, and to promote interpretation of these measures among teachers and administrators, enhancing assessment literacy (Sharratt & Rolheiser, in press). Schools choosing to participate in this program send an administrative/teacher team to a number of District training and follow-up sessions during the year. This Program is credited with creating the groundwork for the effective use of assessment across the YRDSB.

The Ontario Ministry of Education has directed all school districts to select assessment tools for use in their schools. YRDSB had identified a number of specific assessment tools to be used at particular levels in the system. The tools used in elementary schools to assess literacy include PM Benchmarks, the Developmental Reading Assessment (DRA), and Running Records. In addition to these District-wide assessment tools, the EQAO (provincial standardized standardization)
test) results provide grade 3 and 6 Reading, Writing, and Mathematics scores for every elementary school across the province. Overall, teachers and principals are now responsible for submitting data gathered by these assessments to the District on a regular basis, while the District assumes responsibility for analyzing and reporting such data.

The concept of assessment literacy is being promoted not just within the ATP, but in all District work. This initiative establishes both comfort in the use of a range of assessment tools and techniques, and an ability to understand and interpret data, using evidence to guide decisions about practice (Stiggins, 1995).

The second of the District initiatives, Instructional Intelligence (II), began with a group of administrators in 2000; over time, however, school teams consisting of administrators and teachers volunteered to work together on this initiative. The intent of the District was to support teachers in becoming more consciously skilled in using instructional approaches that have a strong research base in terms of their impact on student achievement (Bennett & Rolheiser, 2001). As noted by Sharratt and Rolheiser (in press), “This focus on instruction created greater coherence and connection to our work in assessment literacy…and became a very important stepping-stone for us on our journey to improvement” (p.9). The II initiative included training institutes and demonstration lessons, and follow-up sessions and support provided by the curriculum coordinator and consultant team. Over time teacher leaders have undertaken this important work and contributed to the spread of the instructional ideas (Bennett, Sharratt & Sangster, 2003). Other factors that have also contributed to the sustaining of II work six years later include: a Steering Committee (with multiple stakeholder representation); two OISE/UT preservice cohorts (elementary and secondary) that are located in YRDSB schools; graduate student II research; and, professional development activities specifically focused on developing the instructional leadership of school administrators.

Although early literacy pilot work began in 1999 in YRDSB, it has only been since 2002 that the primary goal in the YRDSB Plan for Continuous Improvement has been literacy. Since that time the District has been focusing its attention on identifying the
challenges in improving literacy and on building the capacity in schools
to make the necessary changes. The manifestation of that focus is a
program called the Literacy Collaborative (LC), the third of the
District initiatives we describe here. This program has evolved from a
grassroots initiative in the late 1990s to what was by 2004 the major
part of the activity of the Curriculum and Instructional Services
Department of the District. Currently all 167 schools in the District are
involved in the LC.

Under the guidance of the Literacy Collaborative each school
must make a three-year commitment to provide release time for a
designated Literacy Teacher (generally a 0.5 position, although some
schools can allocate 0.25 time for this work). The District does not
fund this position, so schools must find the necessary time within
existing staff allocations. Some creative funding with Special Education
Resource Teacher (SERT) positions has enabled schools to establish
this position. As well, elementary schools must also commit to creating
a daily two-hour literacy block for students.

Each school establishes a literacy team, comprising the
Literacy Teacher, the principal, and a SERT; and, depending on
whether it is an elementary or secondary school, the team may also
include the Reading Recovery™ teacher, a division Lead Teacher, the
ESL teacher, and perhaps a classroom teacher as well. This team
attends change/leadership development days and literacy content
sessions offered by the District; attendance at these sessions, which
are supported by the District providing supply teachers to the school, is
mandatory. These District professional development sessions for the
LC are grounded in the belief that leaders across the system need to
not only be skilled and knowledgeable in literacy content, but also to be
skilled and knowledgeable as change leaders. There must also be a
commitment by the team to collect data on the progress being made in
literacy and report this progress in a yearly District culminating event,
the Literacy Learning Fair. Not all of the participating schools have
implemented the program fully. Some schools have a 0.25 Literacy
Teacher position, due to scarce resources or the small size of the
school, while others have not yet implemented the two-hour literacy
block.
In our analysis of the work of these three initiatives in the YRDSB, we have seen how each one of the three initiatives has had an impact on the other two, and that the net effect of all three has been to create stronger capacity at the school and district levels. We now present the findings from our research that led us to this realization and our new model (see Figure 2).

Methodology

For the research described here we developed case studies of four elementary schools in the YRDSB. The four school-level case studies and the cross-case analysis were a means to helping us better understand the phenomenon of capacity-building at the school level and the influence a district might have on this process (Merriam, 1998). We were looking for examples of capacity building and evidence of factors that appeared to affect how, in this case, elementary schools were able to build capacity as part of the Literacy Collaborative.

We have tracked, over two years, the evolution of school capacity in four elementary schools at different stages of implementation of the District-supported LC. This research study documents the process of evolving professional learning, describing the successes and challenges of the schools’ developing capacities within the context of District-level supports. This paper is based on research conducted during the 2003-04 and 2004-05 school years. Data collection methods included conducting individual interviews, forming focus groups, videotaping professional learning experiences, and observing meetings.

Findings

We present the findings of our research on the implementation of the Literacy Collaborative under two headings: Instructional Focus, and Assessment and Evaluation.
**Instructional Focus**

The term ‘instructional focus’ has been used to direct attention on classroom practice. Elmore and Burney (1999) describe their first lesson learned from New York’s District 2 as, “It is about instruction and only about instruction” (p. 239). Their research demonstrates that it is the single-minded focus on instructional practices that brings about improvement in student learning and achievement.

This focus on instruction is central to the design of the Literacy Collaborative. Sustained District-sponsored professional development has helped participants learn about the instructional practices that have the greatest impact on student learning. At the same time, the program has been working to build communities of practice in schools that share this focus on instruction and has encouraged teachers to learn from one another, creating, in effect, professional learning communities where job-embedded learning is the norm. The collaborative nature of the schools increases opportunities for all teachers to focus on quality instruction. Instructional focus has been an explicit dimension of the Literacy Collaborative from its inception.

At Ambleside Public School (pseudonyms are used for all four case schools), the instructional focus was reinforced by the modeling of school leaders. For instance, the principal, vice principal and teacher-librarian all took guided reading groups, demonstrating the importance of these practices for the staff. The sustained nature of the school’s instructional focus also contributed to the strength of the instruction. It was clear that this focus developed over time, and after a number of years, inquiry and attention to quality instructional practices had become critically important to the staff of the school. Indeed, it appeared to become part of the culture of the school.

The changes at Windermere Public School were linked to the instructional leadership style of the principal over an extended period of time. She encouraged staff to be inventive and take risks in finding new ways to help students achieve at higher levels. The establishment of a common language around instruction and literacy, and a consistent approach for building strategies across grade levels, helped create coherence of programs and practices across the school. The staff recognized that the good work done in one class or grade was
Blair Mascall & Carol Rolheiser

contribute to the success in later grades.

Teachers at Windermere credited the Literacy Teachers with making a significant difference to their focus on instruction, through the creation of a shared vision, the development of new knowledge and skills, and the modeling of new practices. Yet teachers acknowledged that they still had work to do; they needed to turn data collection into a continuous practice and to work with new staff to ensure that all teachers were in tune with the changing practices. There continued to be a concern, however, about sustainability of the LC in the District. As one teacher at Windermere said,

I’ll be interested to see how the District continues this and embeds this in instructional practice, long term. Until it’s really who you are and part of your repertoire as an instructor, it’s easy to fall back on old habits when you’re not too sure what to do or you don’t have some form of format or structure to keep pushing everybody forward.

At Green Meadows Public School, staff and students were dealing with the move to a new school in a new community. Staff were chosen for the new school on the basis of their commitment to the importance of literacy in their teaching. This instructional focus was established from the first days of the new school. As staff came to understand how different the present student population was from the population of their previous school, they were better able to understand what it would take to address the challenges.

The staff’s search for ways to redefine their classroom practice in light of the needs of their new student population required that staff spend considerable time identifying student needs and then working on strategies to meet those needs. Green Meadows staff recognized that they would need frequent and targeted data collection; in the first year of the new school, the staff met monthly in their divisions and spent time looking at charts of student achievement, with a particular emphasis on students “at risk.” They started to track student reading levels on a wall chart so that all the staff could see the challenges and the progress that was being made. That action stimulated dialogue among the staff about patterns in the data and the
Pedagogical Synergy: Linking Assessment, Curriculum, and Instruction

differences among reading scores, EQAO tests, and DRA and PM Benchmark results. All of this data collection and dialogue had been carried out in an intentional way, based on the learning from the Literacy Collaborative.

It is important that the District’s focus on instruction is concerned with actual practice. At Jefferson Public School, the development of writing exemplars galvanized the staff to work together on analyzing student work. This process brought coherence to assessment practices, as teachers reached agreement on, for example, “what a Level 3 really looks like.” This collective work encouraged more of the staff to engage in the focus on literacy, as teachers collaborated on instructional planning for a grade and on providing for transitions from one grade to the next. The staff dialogue expanded to encompass instructional and assessment practices across a division and even between divisions.

It is difficult, however, for staff and leadership to sustain a focus on quality instruction in times of change. All our schools experienced such changes during our two years of research. New teachers need to be given the opportunity to learn what long-serving teachers take for granted. New leaders need to take time to learn about the culture of their new school and to find ways to introduce change that does not disrupt the good work that has already been accomplished. Given the frequent changes in leadership that seem to characterize York Region, the District is attempting to develop leadership that can accommodate this mobile workforce.

Assessment and Accountability Focus

The second area explored in our examination of the Literacy Collaborative was the focus on assessment and accountability. While many educators cringe at the word “accountability” we are using this term explicitly here as we connect it to evidence that emerges from assessment practices. Much of the focus on accountability in the YRDSB reflects an effort on the part of the District to build knowledge about assessment and to build accountability across the system by making the use of evaluation data an explicit part of every teacher’s and school’s practice. The use of common assessment tools,
already part of the Assessment Training Program, is also reinforced as part of the Literacy Collaborative. The current focus of the District is on evidence-driven instruction – an attempt to encourage teachers to base their decisions about their classroom practice on the learning and achievement of their students. This practice requires frequent use of a range of assessment tools, a sophisticated interpretation of the results of these assessments, and access to a wide range of pedagogical strategies to respond to the diverse needs of children. This assessment-driven approach has become an integral part of the Literacy Collaborative, in both the change/leadership professional development sessions and in the literacy content professional development sessions with District staff.

In our case schools, all the staff had come to recognize that decisions about classroom practice were to be based on an understanding of the results of their assessment of student learning. Yet, how this played out in practice was not consistent across these schools.

At Ambleside P.S., for example, administration and staff recognized that they did not have a high level of expertise in assessment and were committed to learning more about how to analyze student achievement data. Data were needed to inform the newly prescribed school planning process. Teachers were “on the same page” about evaluation. They used a common language and a common set of approaches for assessing student achievement.

Despite the positive feelings about progress in the school, the provincial test (EQAO) scores had not improved at Ambleside; indeed, they declined somewhat in the second year of our data collection. Staff were disappointed, but started to work on understanding what these results could tell them.

Schools are challenged to manage all of the data that they collect. E-Principal (a District-wide database that gives principals access to all data about individual students) has helped with data management in YRDSB, and a District data warehouse approach is under development. However, it continues to be difficult for schools to deal with the volume of data required by the District.

Dealing with the subjectivity inherent in analyzing data is a challenge commonly identified and discussed by educators in the
District. For example, to compare results across report cards, teachers need to be “on the same assessment page.” Teachers need to know (and agree upon) what a “level 2” looks like. Provincial exemplars are used by many teachers, who also noted that they use them in explaining student achievement to parents. Jefferson P.S. was in the early stages of systematic data collection. How the data might inform teaching and learning had begun to be explored, but classroom teachers had varied levels of understanding about and comfort with analyzing and interpreting data, and with utilizing the data to inform practice. The school complied with District data collection requirements, but the results were unevenly used. There was also some difficulty in interpreting the differences between data sets – for instance, EQAO and PM Benchmarks showed quite different trends in achievement. Teachers struggled with synthesizing the information from these various sources.

Teachers and administrators at Green Meadows P.S. were turning to assessment data to help them understand their new student population. The differences between data sets merely fueled further investigation, and there was an ongoing effort to look for ways in which these data could inform instruction. Teachers and the principal actively sought to find connections and patterns between data, consistently comparing their own locally developed assessment results, the District PM Benchmark and DRA data, and the provincial EQAO findings. There appeared to be a genuine spirit of enquiry at Green Meadows, which encouraged the staff to use a variety of data sources to understand more clearly the needs of their students.

Windermere P.S. was populated with a staff of action researchers, providing evidence of the role that action research has played in YRDSB. Their conversations were filled with questions, as they explored ways to enhance student learning and contribute to enhancing the learning community. Teachers were using the language of assessment as part of these conversations. Indeed, the language was becoming part of the children’s discussions as well. During a session where students were putting samples into their writing journals, one student commented to another, “I don’t think that’s a level 3 letter any more, I think it’s a level 4.” As the teacher noted, “…all that language about assessment. I wouldn’t have heard that in a group a
year ago. So it’s not just at the teacher level, students are understanding their own learning.” Assessment literacy at Windermere appeared to be pervasive.

The accountability at Windermere had become shared. Use of such mechanisms as student tracking boards encouraged staff conversation about the achievement levels of individual students and put faces on the data. Teachers were analyzing the data saying, “Here’s a child at risk. What do we need to do? How can we tailor the instruction to meet her needs?” This shared concern produced a broader sense of accountability across the staff.

Parents on the Windermere School Council were asking for feedback on how the resources they helped provide were contributing to student achievement. Such observations were indicative of an increased level of public accountability that would not have been present even a few years earlier. The degree of comfort felt by the staff was evident in the high level of trust they showed each other. People were not trying to hide results. Rather, they presented results to each other and discussed how one might interpret the findings and what a teacher might do in light of this interpretation. Staff at Windermere expressed gratitude for the expert support they received from District personnel in managing and interpreting data. This support had made a difference in the ability of the teachers to understand what the data were saying, and to use that understanding to design future learning opportunities.

**Pedagogical Synergy**

We have been describing the three components of instructional quality relating to Newmann King and Youngs’ (2000) model of school capacity: curriculum, instruction, and assessment. Although one might take as self-evident that these three components are essential for enhancing student achievement, in practice what often happens is that district or school initiatives related to these dimensions operate independently or compete with each other, producing either fragmentation or overload for educators.

In York Region, three programs (the Assessment Training...
Pedagogical Synergy: Linking Assessment, Curriculum, and Instruction

Program, the Literacy Collaborative, and Instructional Intelligence) appear to have come together in the minds of educators in a complementary way, in fact, seeming to create what we have termed “pedagogical synergy”. Based on our research and conceptualization, we define pedagogical synergy as a dynamic enhancement of understanding, strengthening of practice, and development of coherence. This synergy occurs when curriculum, instruction, and assessment link with and support one another, rather than operating as separate or conflicting entities.

In the early days of our previous research (Mascall, Fullan & Rolheiser, 2001), these programs were seen as discrete and independent; and for some teachers, they were competing programs, demanding scarce resources (time, money, materials). Over the years of our research, we have noted how educators now are seeing ways these two programs fit together and support each other. In discussions about the Literacy Collaborative, teachers and principals told us of the work they were doing to improve classroom practice, informed by their work in the Assessment Training Program and Instructional Intelligence. They also spoke of how their literacy assessment has changed as a result of the Assessment Training Program. They recognized the common learning approaches and the common language associated with these programs. For example, the use of action research and other job-embedded professional development strategies have been integral aspects of all of these district-supported initiatives. Teachers clearly recognized that inquiry and continuous learning of educators was valued by the District.

This evolution has not happened in a casual way. The District has taken a strategic approach in selecting a few key foci and then targeting resources and aligning policies to ensure that these become central in the work of all educators throughout the District. In the District approach to leadership development, for example, it is clear that the expectations for administrators in the District encompass a focus on the targeted and selected District initiatives, including the Literacy Collaborative. New leaders are required to understand and support this relationship and to see it in practice in their schools.

From the perspective of our research beyond this particular study, we have been able to monitor the evolution of this strategic
approach and have seen in our latest study that the engagement of Board members, the ongoing discussions among senior management, and the deliberate and sustained professional development with teachers and administrators have resulted in reduced fragmentation and overload among school staffs. At the same time, it has reinforced the importance of instructional quality and enhanced school capacity.

**A New Model**

From this research, we can see that our initial model has the right pieces: we believe we have identified key dimensions in both school capacity and district capacity. These dimensions remain the same. However, we would suggest that putting them in boxes, in a linear order (as we did in Figure 1), suggests a vertical (hierarchical) approach to the building of capacity that we feel might be misleading. We propose a new conceptualization of our model, described in Figure 2.

Our research in York Region suggests that these items are more powerfully conceptualized horizontally, with movement back and forth between the District and the schools. Our case schools see their work as being done in collaboration with the District, with both partners clearly focused on improving the literacy results for all students. The role of the District is still pivotal in terms of setting an overall vision, providing appropriate kinds of support, and making changes in policies and procedures to align with their literacy vision. However, school teams are empowered to implement their contextually unique approaches to literacy based on the District’s guidelines, and this local learning yields important insights that subsequently influence the District’s ongoing work.

Desired results can occur at both the district and school levels when there are horizontal and reciprocal strategies for building capacity and increasing coherence. During the two years of our study such strategies became more evident and were more explicitly applied. With a united and coherent emphasis on curriculum (in this case, literacy as the focus), instruction, and assessment, *pedagogical synergy* is created. Our model describes a collaborative culture in which schools and districts support each other, rather than a counter-productive struggle to assert top-down or bottom-up ownership.
Pedagogical Synergy: Linking Assessment, Curriculum, and Instruction

Schools and the district are united in a focus on results for students. The reciprocal nature of the work in York Region means that vision and direction provided by the District influences school actions, while school actions and insights also influence subsequent District understandings and actions. This continuous flow of strategies builds synergy over time as curriculum, instruction and assessment become more highly integrated. The result is greater system-wide capacity, and a coherence in teaching and learning which supports both district
and school educators.

This new model stresses the interconnectedness of the variables (they all seem to influence each other), and the horizontal (non-hierarchical) nature of the relationships among the educators. It describes an approach to education that has resulted in greater coherence across the system.

We see the experience and growth of York Region over the two years of our research as positive outcomes. Recent research by the District (Sharratt & Fullan, in press) concludes that schools which are dedicated to implementing the Literacy Collaborative model fully over time can be expected to out-perform schools that are not. Evidence also indicates that over 80% of Grade one students in YRDSB are reading with fluency and comprehension at Level 16 (the District standard), an increase from 51% when the Director of YRDSB first made his commitment to a literacy focus in 1998 (Sharratt & Fullan, in press). As well, other research (Sharratt & Rolheiser, in press) indicates that the strategic work of the District is resulting in important gains in student learning over time. The successful EQAO (the provincial standardized test) results in literacy and mathematics over the last five years are attributable, at least in part, to the strategic work of the District and the focus on and weaving together of sound assessment and instructional practices (Sharratt & Rolheiser, in press). Certainly the evidence we saw in our case schools about building collective capacity would support such findings.

It will be important for York Region to sustain its efforts and remain focused on knowledge building about literacy. The pressure to address competing issues will be hard to resist, but we believe this model needs to be supported for the long term. In the evolution of the work in York Region to build capacity between districts, the District may worry that it will lose the strength of the internal focus it has cultivated. The staff should guard against overload and the distractions of competing policies and demands. The recent lateral shift to support work between districts is important outreach and has the potential to contribute to knowledge sharing; however, it will be important to simultaneously sustain the internal focus and energy that was evident in York Region schools during our research.
References


Bennett, B., Sharratt L., & Sangster S. (2003). A focus on Instructional Intelligence...one and one-half years into a five-year journey. Paper presented at meeting of ICSEI, Sydney, Australia.


From Performance-Based To Inquiry-Based Accountability

Dr. Sonia Ben Jaafar (Ph.D.)
Ontario Institute for Studies in Education

Abstract

This paper argues Canada has a unique approach to large-scale assessment and accountability that is not reflected in the current literature. I explore its construction by drawing on the results of two pan-Canadian SSHRC-funded studies and reflecting on Canada's philosophical identity that distinguishes it from other nations. I conclude with a call for more Canadian-based theoretical and empirical research to support contextually relevant interpretations in future studies.

Large-scale assessment for accountability has become an inherent part of educational reform over the last decade (Popham, 1999; Ranson, 2003). Although Canadian educational reform adheres to this global trend, I argue it has a unique approach that is not reflected in the current literature. This paper explores the construction of assessment and accountability in Canadian education by drawing on the results of two pan-Canadian SSHRC-funded studies. These studies

Dr. Sonia Ben Jaafar recently completed her doctoral degree in the Department of Theory and Policy Studies at the Ontario Institute for Studies in Education of the University of Toronto (OISE/UT). Her research interests include system evaluations, public accountability policies and practices, organizational leadership, and international comparisons. She intends to continue pursuing these interests on a more international scale.
focused on the social consequences of test use and the implications that grow out of that use (Messick, 1989). In an attempt to better interpret and understand the findings, I explored Canada’s philosophical identity that distinguishes it from other nations as one built on a foundation of compromises (Saul, 1998).

The idea for this paper was inspired by the intersection of a series of academic and political activities that were part of my graduate education. In November 2005, I attended a two-day forum, Unlocking our Children’s Potential: Literacy and Numeracy as a Foundation, jointly held by The Council of Ministers of Education, Canada (CMEC) and the Literacy and Numeracy Secretariat at the Ontario Ministry of Education. At that time, I was struggling to understand the relationship between performance-based accountability policies and the responsive practices in Canadian schools. The speeches, presentations, and conversations at the forum influenced my appreciation for Canadian accountability in public education. The discourse conflated assessment, learning, and achievement with accountability, which reflects the dominant position of large-scale student assessments as a mechanism for educational accountability. This paper presents my reflections on educational accountability in Canada given my recent interactions with provincial policymakers and my involvement with two recent pan-Canadian studies that examine educational accountability policies.

This paper is divided into three sections. First, I will introduce the two studies that raised questions about the unique nature of Canadian educational accountability. Second, I will present a discussion intended to raise awareness of the unique nature of Canadian identity and how it relates to educational accountability. Finally, I highlight challenges for Canadian scholarship in this field that extends to generalizations about valuing Canadian knowledge. I also peppered contemplative commentaries throughout the paper to highlight how my perspective has evolved.

**A Pan-Canadian Look at Educational Accountability**

I was involved in two studies that examined how Canadian public
policy in education is operationalized (Ben Jaafar & Anderson, in press; Ben Jaafar & Earl, in press). By virtue of their seniority in the field, the leadership and supervision of Drs. Lorna Earl and Stephen Anderson ensured a historical perspective of Canadian policy informed the work. Moreover, the results of each study informed the interpretation of the other, illustrating the importance of promoting Canadian scholarship. Being privy to this complementary process was an important foundation for the inception of this paper long before its development.

The first study is embedded in a five-year pan-Canadian investigation of the work lives of public school teachers and principals and the policy and social conditions that are influencing their work. We conducted a logical policy analysis focused on the internal and horizontal consistency (Pal, 1997) of educational accountability policies in Canada’s thirteen jurisdictions. Narratives addressing educational policy change from 1990 to 2003 at the government level for each of the 13 provinces and territories served as the database for a pan-Canadian look at policy trends associated with accountability. The analysis employed Spencer’s (2004) conceptual framework describing and differentiating Economic-Bureaucratic Accountability (EBA) and Ethical-Professional Accountability (EPA).

We concluded that Canadian educational accountability is best described as a hybrid-model of results-oriented EBA and process-oriented EPA. Although the EBA approach currently dominates the accountability policy discourse, our analysis illuminated how the two accountability models co-exist with a notable partition in accountability for education management and professional services (Ben Jaafar & Anderson, in press). In simplistic terms, the EBA model is employed for the management functions of schools, whereas the EPA model is employed for the teaching or the process of educating students. At the juncture where the management and service of education interact, the 13-year trend indicated a tension between the two models that suggested an overall balance over time between the two models.

The second study is part of a three-year examination of the relationship between performance-based accountability models and the responsive practices of secondary school leaders across Canada. The phased mixed-method research was conducted over two years.
From Performance-Based To Inquiry-Based Accountability

(2004-2006). The first phase examined the performance-based accountability policies of the provinces and territories across Canada. Rich case descriptions were used for a cross-case pattern analysis to capture the similarities and characterize Canadian PBA model types (Ben Jaafar & Earl, in press). The second phase of the research examined the PBA-related practices of secondary school administrators in two provinces that were identified as housing distinct PBA models, Alberta and Ontario. The secondary school administrators in a randomly selected sample of districts in each province were sent surveys asking questions about their school practices associated with assessment and accountability. The data was subject to comparative analytical procedures that tested for differences in the administrators’ practices. Finally, the third phase examined the relationships between the provincial PBA policies and the school-level practices. A pattern analysis was conducted searching for relationships between school practices and provincial PBA policies.

The findings in this study indicate that secondary school administrators can be categorized into three types: Consequence-prone, Consequence-averse, and Moderate school leaders. Most administrators in both provinces are Moderates, which means that they respond to the provincial assessment results with collaborative and bureaucratic practices in their schools. There were no relationships detected between any school or PBA policy dimensions and the practices of school leaders. The only variable that related to their practices was a self-declared belief in the utility of the provincial assessment and accountability system for schools (Ben Jaafar, 2006).

Raising Questions and Raising Canadian Knowledge

There was an obscure attribute to the findings in both these studies. They did not corroborate the other studies in the literature. For example, I found that the beliefs of school leaders were the key factor determining how they behaved in schools (Ben Jaafar, 2006). This was in direct contradiction with the findings of Ladd and Zelli (2002; 2003), who concluded principals in North Carolina aligned their practices with
the policies irrespective of their beliefs. The distinction is most likely due to the disparity in the use of high-stakes between any of the Canadian jurisdictions and North Carolina. The problem is that the current literature addressing assessment and accountability generally assumes higher stakes than those in Canadian models. This raised the issue of how to draw on the literature to understand the Canadian findings?

I was struggling to make sense of the findings in relation to the current literature on performance-based accountability, when a U.S. scholar who recently arrived in Ontario shared an important observation. He claimed that Canadian educational accountability is superficial and that we do not have “real” accountability mechanisms for schools. Given the work with which I was involved, and the difficulties I was having in situating my findings in current literature, his comment gave me pause to reflect on the notion that my studies were socially constructed works of fictions.

The scholar had articulated his perception of educational accountability in a sincere effort to better appreciate the Canadian context. As is the case with anyone attempting to make sense of a new context, he drew on his established knowledge and experience. His frame of reference was an educational accountability model constructed by a reaction to A Nation at Risk (National Commission on Excellence in Education, 1983) to the current reality of No Child Left Behind (U.S. Department of Education, 2003). His innocent observation illustrates the danger of “rhetoric spillover” from our neighbours which imposes inappropriate ways of understanding because they do not resonate with the Canadian approach to public policy (Ungerleider, 2003). I submit that the same argument Ungerleider made for rhetoric spillover can be made for the academic knowledge we value. Given this cultural bias inherent in the academic knowledge valued on this topic, I sought out different kinds of Canadian knowledge to understand educational accountability rather than what is in the traditional academic venues.

**Accountability and Dynamic Equilibrium**

The Canadian story of educational accountability takes many shapes.
At the CMEC/LNS forum, a running theme was that all students should reach high standards; a theme that would be prevalent in many nations addressing education (See OECD PISA participation list). However, as educational leaders from the provinces described, referenced, reflected on, and engaged with the idea of educational accountability, they focused on equitable opportunities to learn. There was an assumed Canadian value of social justice embedded in the conversations. This value translates into responsibility for educating individuals for citizenship in a school system, rather than a responsibility for an education system.

This sense was reflected in the forum opening remarks made by The Honourable Gerard Kennedy, the Ontario Minister of Education at the time. He congratulated the assembly of systems thinkers who influence education and appreciate their mandate. He recognized that the participants were “prepared to take responsibility for whether their ideas work and can affect the lives of children” (Kennedy, 2005). He couched the claim by identifying the root promise of education as supporting students to meet their potential at the earliest age possible to meet the essential purpose of education: To expand citizenship. Kennedy left the group with the idea that social cohesion in modern times requires a concerted effort to include all individuals in society.

The significant position of schools in promoting Canadian citizenship for social cohesion becomes apparent when democracy is thought of as an expression of shared knowledge that comprises a society’s common sense (Saul, 2001, p.28). Public schools are the institutions where the foundation of shared knowledge is learned and established across all people in society irrespective of socio-economic, cultural, linguistic, or religious status. Saul (2001) argues that schools are the last public institution where the majority of the population conglomerates. As that last meeting place, it is a requisite for public education to respect the diversity of our complex society.

In a complex society, social relations do not fall neatly into tight definitions that demarcate who should know what. The complexity of Canadian society is an organic acknowledgement of the diversity of its population and geography that have different requirements. Canada has never embraced the illusionary mythology of unity in its conception
of social cohesion. Our social cohesion moves “beyond the narrow, unitary concept of nineteenth-century nationalism towards an equilibrium between groups and cultures” (Saul, 1998, p.241).

This vision is paradox to the European and American stated ideals of the nation-state with a solitary unifying identity. The image of a single identity is one that is theoretically viable, but collapses in practice. The attempt of imposing one in Canada as a spillover from our neighbours would merely convert our celebrated complexity into division (Saul, 1998). Theory and practice are tightly coupled in Canada where leaders make policies representing theoretical ideas into an isuma practice. This tight coupling causes a natural discomfort that requires concerted individual and collective effort to challenge implicit values and beliefs. Incongruent ideas and practices give rise to tensions and conflicts that require a tolerance for ambiguity. The findings in our pan-Canadian examination of educational accountability illustrates how this tension gets manifested in policy over thirteen years (Ben Jaafar & Anderson, in press).

In reflecting on Canadian values, I argue the discomfort is actually constructive and that the negative connotation associated with the term “tension” is a deliberate outcome of literature promoting the ideal of uniformity as a valued asset (See Blackmore, 1988; Datnow & Castellano, 2001; Pal, 1997). Another way of conceptualizing this “tension” is as a dynamic equilibrium between incongruent, but not incompatible ideas and practices. An equilibrium based on attending to the rights of the individual and society is especially important when a society is composed of a conglomerate of minorities. Monolithic entities have not existed in Canada since the birth of the nation (Saul, 1998).

This equilibrium, which Canada has purposefully constructed and maintained since its inception as a country, constitutes the philosophy of its national identity (Saul, 1998). For Canada, the shared knowledge established as our common sense taught in schools necessarily grows and evolves out of a dialogue between the multiple groups of its population. Valuing multiple identities insists on devolution of responsibility that demands educational leaders engage in the reflective thinking necessary for improving the promotion of shared knowledge. Educational accountability in support of this purpose needs
to accommodate local flexibility as a reflection of the nation’s identity.

When moral imperative is coupled with transparent data, it creates the appropriate change pressure that motivates educational leaders to promote an inclusive society (Fullan, 2005). The responsibility of the local educational leader is to link the local needs to the larger societal ones. In essence, they act as the fulcrum that levers societal equilibrium in schools. Aspiring for equilibrium recognizes complexity and acknowledges the process towards equilibrium is all we can achieve because a perfect balance cannot exist (Saul, 2001).

This approach to education embeds Canadian social responsibility as a foundation of our accountability. It constructs large-scale assessment results as information educators use to know what to do given the local situation and the demands of the broader context. It moves the focus from a view of the results as the final measure of success, to recognizing they are an entrée to identifying where attention should be focused for the appropriate differential provision of opportunities to learn across situations (Betts & Danenberg, 2002; Earl, 1999).

This structures educational accountability as a readiness to take responsibility for actions that bring about greater equilibrium as the Canadian moral purpose of education. So that we consciously act in direct objection with “mid-nineteenth century moralism in which those children without must either be left behind or be humiliated by entering a category deserving of charity” (Saul, 2001, p.60). It means that we approach accountability through questions not answers and that not all stories are always stories of improvement, but they should all be stories of learning (Kaser & Halbert, 2005). This approach dissolves punitive consequences, because “failures” to meet expectations are perceived as fertile ground for discovering an imbalance in educational approaches that informs future improvement. It is an inquiry-based accountability where inquiry is the intelligent uncertainty of assessment-literate educational leaders (Kaser & Halbert, 2005).

Saul (2001) describes the pursuit of equilibrium as a manifestation of compassion rooted in an appreciation of the uncertainty necessary to understand others. This compassion requires enough comfort with ambiguity to construct shared knowledge across
Inquiry-based accountability starts with the questions that create anticipation for results, which can then be purposefully examined to indicate appropriate changes in process. It constructs the results of assessment programs as early warning signs that allow for the anticipation of problems. It is not reactionary with an eye for deficits in the system, but considers where improvements can occur and how that can happen. On the surface, this accountability may look the same as PBA. The difference is that IBA does not focus on the final numbers because they are only measures whose beneficial quality is to direct attention for responsible actions.

For the results of large-scale assessments to operate as mechanisms for school improvement, educators require certain capacities (Carnoy, Elmore, & Siskin, 2003; Earl & Katz, 2003, 2005, 2006; Elmore, 2005; Ingram, Seashore Louis, & Schroeder, 2004). High-stakes punitive PBA structures like NCLB inhibit the development of those capacities (Fullan, 2005). PBA is synonymous with consequences in the literature (Gayler, Chudowsky, Kober, & Hamilton, 2003; Pearson, Calfee, Walker Webb, & Fleischer, 2002; Stecher, 2002). Yet, the promotion of consequences attached to assessment results is insufficient to understand and change educational processes (Mintrop, 2004).

IBA explicitly positions educators so that they take responsibility for making sense of the large-scale assessment results. Educators who work within an IBA framework necessarily build their capacity to examine data and educational processes to improve student learning. The embedded assumption in IBA that differentiates it from PBA is a belief that the investment in educational processes will yield increased student achievement, as opposed to a belief that any means can be applied to increase achievement (See Cable News Network (CNN), 2005; CNN, 2005; Valencia, Valenzuela, Sloan, & Foley, 2004). This distinction helped me think through some of the issues that surfaced when I was interpreting the findings of the two pan-Canadian studies. In the Canadian context, the social consequences of accountability for educators are markedly different from what is predominantly described and discussed in the literature. The PBA systems embedded in the literature promote the compliance of educators through coercive mechanisms, whereas in IBA, changes in
practice are promoted through capacity building exercises. The practical result of this difference is illustrated by the reaction of school leadership. Again, in North Carolina (PBA), principals act in conflict with their beliefs to increase student achievement (Ladd Helen F & Zelli, 2002), whereas in Ontario and Alberta (IBA), principals need to be attuned to their beliefs and those of the community they serve to responsibly make sense of the results and respond appropriately (Ben Jaafar, under review). The implication of the Canadian approach is that it makes room for local values and beliefs.

**When We Don’t Value Ourselves**

Both studies introduced at the beginning of this paper were from Canadian datasets. It is my contention that if this is the case, then the interpretations should respectfully reflect the cultural context of that jurisdiction. This idea may seem so obvious that the reader may think I am being pedantic in reiterating this point. But, established academic protocols demand that data interpretations are substantiated by current literatures. However, when the literature is dominated by a single cultural perspective, interpretations become skewed and it is possible to lose insights that could further our shared knowledge. Canadian research is at particular risk because of our unique identity.

I submit that the flexibility inherent in the Canadian model of inquiry-based accountability reflects the idea of Canada’s aspiration as a nation: A nation aspiring to harmonize a complex array of differences through compromises. Practically, this is accomplished through policies that privilege concessions that respect the identities of local communities. Canadian educational accountability respectfully acquiesces to the values of individual communities. It relies on the commitment of local leadership to understand the local values and use them in conjunction with the results of large-scale assessment to determine the best ways to support the school given the purpose of promoting equilibrium for social cohesion.

Saul (1998) offers encouragement and warning about the global perception of the flexibility embedded in how our nation functions:
The middle way is neither soft nor easy. It is the most difficult of roads, because it is the most fragile and is exposed to easy attack. The ways of ideology and absolute answers, the monolithic view of the nation-state, the dominant centralized view of culture - all of these are filled with bravado. The corporatism which so dominates today has nothing but contempt for a complex middle road. But then corporatism is the logical partner of simplistic nationalism. (Saul, 1998, p.113)

His description highlights the vulnerability of consistently working towards sincere inclusion when building our socially cohesive fabric. The global proliferation of performance-based accountability illustrates how Canada’s approach is lost in the discourse, and consequently devalued. It is why the U.S. scholar I introduced at the beginning of this paper did not even recognize that Canada had accountability measures for public education.

IBA is markedly different than PBA, which dominates the current literature. PBA constructs the results of large-scale assessment as the end of education. It is considered a measure of individual success with an implicit blame on the system for poor results. IBA constructs the results to serve as an entrée to conversations about the opportunities to learn that exist in schools for students of diverse backgrounds in different communities. Implicit in PBA is the importance of blame rooted in a deficit model searching for lack of conformity to an ideological profile. Punitive measures are appropriate in this model because consequences can be individually attached to those who do not conform. This approach is operational in nation-states ideologies where there is a renowned solitary identity, such as in Europe and America. However, the model that fits Canadian society needs to construct responsibility as a progressive endeavour that acknowledges complexity and works towards an equilibrium.

Hess (2002) differentiates the coercive theory on which high-stakes measures are premised from standards-based reforms when he claims “gentler, less threatening standards-based approaches seek to improve schools through informal social pressures and by using tests as diagnostic devices” (p.70). This is compatible with Elmore’s school-
localized description of internal accountability where educators examine the results of large-scale assessments in relation to school processes (Carnoy et al., 2003; Elmore & Fuhrman, 2001). Like these others, Mintrop (2004) offers exceptional cases where external stakes can work when internalized into inquiry processes that promote professional learning. Interestingly, researchers working in environments that impose a monolithic view of a nation-state identify these “exceptional cases,” which are essentially localized examples of IBA within PBA systems.

Canadian IBA is intended to reflect compromises across diverse perspectives as something more than a desirable nicety. It has been a necessary element of Canada’s survival since 1867. Compromise in Canada is part of our identity, and because of that, our mechanisms for public policy reflects a local flexibility that promotes dialogues of shared knowledge between groups of minorities, who together represent our population.

There are a number of scholars who examine the use of large-scale assessment results in a way that can be supportive to school improvement. They are working in a context where PBA overshadows the shared knowledge because they employ interpretive explanations that draw from a literature that does not represent our cultural context. The work is already challenging without the added obscurity of insufficient Canadian theoretical explanations for the phenomenon. It is unsatisfactory to adopt U.S. and European theoretical frames of understanding to interpret Canadian data, or vice-versa. We may and should welcome the consideration and adaptation of all frames of understanding in research. But when it comes to constructing knowledge based on our findings, the explanations need to reflect the values that founded the nation and are still a way of life. It is, after all, common sense.

1 L’évolution du personnel de l’enseignement des écoles au Canada is a SSHRC Major Collaborative Research Initiative. The project direction is situated at the University of Montreal. The project team includes researchers from 10 universities and Statistics Canada.

2 Performance-based accountability in policy and practice: implementation and
influence of secondary school assessment policies across Canada is supported by The Social Sciences and Humanities Research Council of Canada. The project team is from the International Centre for Educational Change at OISE/UT.

3 Isuma is an Inuktitut term that “encompasses a sense of responsibility toward the community” (ISUMA: Canadian Journal of Policy Research, 2003).

References


Résumé

Cet article traite des représentations du personnel enseignant à propos des plans d’enseignement individualisés (PEI) destinés aux enfants en difficulté. L’intention est de cerner le lien entre le point de vue de ces intervenants et certaines de leurs caractéristiques personnelles et professionnelles. Une enquête a été menée auprès d’un échantillon du personnel enseignant francophone de l’Ontario oeuvrant dans trois cycles d’enseignement différents. Les résultats laissent entrevoir, entre autres, que les représentations que se font les enseignants des PEI sont influencées par leur ancrage professionnel.

Introduction

Enfance en Difficulté

Plan est d’abord précédée d’une évaluation formelle de l’élève par un comité d’identification de placement et de révision du conseil scolaire. Les informations pertinentes au sujet de l’élève sont par la suite transmises à l’équipe qui prépare le PEI de l’élève. Cette équipe se compose de personnes qui connaissent et qui travaillent avec lui, notamment le personnel enseignant et toutes autres personnes qui ont la responsabilité de répondre aux besoins particuliers de l’élève en question.


Problématique

Depuis que la préparation d’un PEI est devenue obligatoire dans le système d’éducation ontarien, il y a eu énormément de controverse quant à son utilité pour l’enseignant. Dans les milieux scolaires, il est « commun » de qualifier cette pratique de « cancer pédagogique » ou de « tâche administrative qui contribue très peu à l’éducation des
enfants en difficulté ». Bien que cet état d’esprit soit documenté aux États-Unis (Bateman, 1995), il reste à mieux comprendre ses diverses dimensions en Ontario. Nous voulons donc apporter un éclairage sur cette question en vérifiant les représentations du personnel enseignant à l’égard du PEI. Quels facteurs influencent les représentations des enseignants? Qu’en est-il des stratégies proposées? Répondons-nous aux besoins des élèves en difficulté en utilisant un PEI ? Ce sont là autant de questions auxquelles nous avons tenté d’apporter quelques éléments de réponses dans la présente recherche. Notre hypothèse est que les représentations que se font les enseignants du PEI sont davantage influencées par leur expérience scolaire et le niveau d’enseignement que par des caractéristiques personnelles comme l’âge ou le sexe. Ainsi, le fait d’œuvrer auprès d’élèves d’âges variés et à des niveaux de scolarité différents agirait plus directement sur les représentations que les enseignants se font du PEI.

**Recension des écrits**


Dans la recherche comme dans la pratique, nous vivons toujours dans la période « infantile » des PEI (Bateson, 1995). Aux


L’accent mis sur les aspects pédagogiques du PEI est important puisqu’il nous renseigne sur la recherche des voies visant à mieux répondre aux besoins des élèves. Toutefois, il laisse dans l’ombre les attitudes et les conceptions des enseignants qui sont loin d’être négligeables dans la mise en œuvre des PEI. Cet aspect est peu traité dans les études. La recherche de Thomson et Rowan (1995), qui s’intéressent en partie sur les perceptions et représentations d’enseignants et de parents néo-zélandais sur le lien entre le PEI et l’éducation des enfants à besoins spéciaux, est l’une des rares études dans le domaine.

La recension des écrits nous laisse croire que les connaissances disponibles au sujet du PEI résultent parfois d’études...
empiriques, alors qu’en réalité elles demeurent encore largement théoriques. Ce sont les descriptions qui dominent à partir de points de vue traduisant la vision idéale que les auteurs ont des PEI et de leur utilisation auprès des élèves, ainsi que le montre, par exemple, l’étude historique de Smith (1990) dans le contexte américain. Par ailleurs, il apparaît que la plupart des études sur le PEI sont dépourvues d’assises théoriques explicites, ce qui limite généralement la fécondité des résultats.

**Cadre théorique**

L’hypothèse de cette recherche est formulée sur la base de la théorie des représentations sociales (Moscovici, 1961, 1989). Notre perspective est d’établir le rapport des enseignants au PEI comme le résultat non pas d’une expérience subjective mais d’une construction idéologique collective du groupe professionnel du personnel enseignant. La théorie des représentations sociales s’avère éclairante à cet égard. Par exemple, Jodelet (1989) conçoit une représentation sociale comme étant une forme de connaissance socialement élaborée et partagée, ayant une visée pratique et concourant à la construction d’une réalité commune à un ensemble social donné. Si elles influencent les conduites et la communication des groupes humains et de leurs membres, les représentations sont elles-mêmes façonnées par les rôles et l’expérience des acteurs sociaux (Doise et Palmonari, 1986). Parmi les processus qui participent à la formation et au fonctionnement des représentations sociales, les théoriciens se réfèrent, entre autres, à la notion d’« ancrage ». Cet ancrage permet de comprendre les catégories nouvelles porteurs d’une information plus ancienne, établies et partagées au sein du groupe. Le PEI, comme mesure pédagogique et légale relativement nouvelle dans le champ scolaire, pourrait bien se comprendre dans le rapport que le personnel enseignant entretient avec lui, à la lumière des représentations sociales.

**Méthodologie**

Les données de l’étude proviennent entièrement d’une enquête
par questionnaire auprès du personnel enseignant d’écoles francophonontariennes. Deux étapes ont marqué la conception de l’instrument de recherche. La première a consisté à sonder le pouls d’enseignants au sujet des PEI afin de bien cibler les questions de recherche. Pour ce faire, quatre groupes focus ont été formés: deux dans la région du moyen nord de l’Ontario, et deux dans la région de l’est de la province. Le but était d’avoir un échange libre avec les participants et participantes pour circonscrire les représentations les plus souvent évoquées dans les milieux scolaires au sujet du contenu des PEI. Au total, vingt membres du personnel enseignant des régions mentionnées ont participé à l’exercice.

La deuxième étape a mené à l’élaboration d’un questionnaire faisant appel à une échelle Likert avec des niveaux allant de 1 à 5: le chiffre « 1 » signifie que vous ne vous adonnez jamais à l’activité qui est désignée par l’énoncé; le chiffre « 5 » signifie que vous y adonnez très souvent; les chiffres « 2 » à « 4 » correspondent à des positions intermédiaires. Le questionnaire a été construit en tenant compte des points de vue exprimés par les enseignants lors des groupes focus. Il comprend cinq variables à expliquer: 1) le lien entre l’information contenue dans le PEI et les besoins des élèves, 2) le caractère pratique des stratégies d’enseignements proposées dans le PEI, 3) la fonction attribuée au PEI pour mieux connaître les élèves, 4) l’utilité pédagogique du PEI et 5) la pertinence des attentes des élèves dans le PEI.

Une question ouverte permettait aux répondants de partager d’autres commentaires au sujet des PEI. Les questions d’âge, sexe, années d’expérience, niveau d’enseignement et le nombre de cours suivis en enfance en difficulté ont servi de variables indépendantes.

Échantillon

Parmi les douze conseils scolaires francophones de la province de l’Ontario, huit ont accepté de participer à cette recherche qui, aux dires de plusieurs directeurs suscitait énormément d’intérêt dans les milieux scolaires. Cinquante questionnaires ont été acheminés directement aux personnes ressources de l’enfance en difficulté de
Yves Gauthier

chaque conseil, qui les ont ensuite distribués au personnel enseignant de différentes écoles. En tout, des 400 questionnaires distribués, nous en avons reçu 329 pour un taux de participation de 82,25%, ce qui témoigne de l’énorme intérêt porté à l’égard des PEI dans les écoles françaises de l’Ontario.

Les répondants proviennent de tous les coins de la province puisque huit conseils scolaires ont participé à la recherche. On compte 251 femmes (76,3 %) par rapport à 78 hommes (21,6%) et l’âge moyen est de 37,4% ans. De ce nombre, 143 participants enseignent au primaire-moyen (47,5 pour cent), 79 au primaire-intermédiaire (23,9 pour cent) et 86 (28,6 pour cent) sont au niveau intermédiaire-secondaire. En moyenne, le personnel enseignant a 12,14 années d’expérience. De plus, 167 (50,8%) indiquent ne jamais avoir suivi de cours de qualification additionnelle en enfance en difficulté, 64 ont suivi la partie 1(19,5 pour cent), 20 (6,1%) et 78 (23,7 pour cent) ont suivi la partie 3; c’est-à-dire qu’ils sont spécialistes dans ce domaine.

Résultats

Les tableaux 1 et 2 présentent des données de base au sujet de l’utilisation des PEI par le personnel enseignant qui a participé à cette recherche. Cette distribution de fréquences présente également les données des répondants pour toutes les questions du questionnaire. Les tableaux nous permettent ainsi d’avoir une vue d’ensemble sur les opinions du personnel enseignant concernant les PEI. Selon l’échelle Likert que nous avons utilisée, le « 1 » signifie que les répondants et répondantes ne s’adonnent jamais à l’activité désignée, tandis que le « 5 » indique qu’ils s’y adonnent souvent.

Ces tableaux tracent un aperçu des opinions du personnel enseignant en lien avec la question des plans d’enseignement individualisés. Nous notons, par exemple, que les répondants n’entretiennent pas d’opinions très positives au sujet du PEI en tant qu’outil pédagogique (X= 3,65 selon l’échelle Likert). Dans le même ordre d’idées on remarque que les répondants ont tendance à être d’accord avec l’énoncé selon lequel ils consultent les PEI pour mieux connaître les élèves (X=4,00).
En examinant les troisième et quatrième tableaux, on peut connaître le nombre et le pourcentage des répondants qui se sont prononcés sur les douze énoncés ainsi que leur position sur l’échelle de Likert. Si on fait fi du nombre de répondants, on remarque que ceux-ci accordent une très bonne note à l’ensemble des variables à expliquer. Seule la dernière au sujet du système éducatif ontarien en regard aux besoins des élèves, fait l’objet d’une forte dépréciation. À l’autre extrême, les enseignants sont nombreux à dire qu’ils consultent les PEI, que l’information que ces plans contiennent reflète le besoin des élèves, que les attentes sont pertinentes et qu’ils s’en servent. Mais ils sont tout aussi nombreux à dire que les normes du PEI devraient faire l’objet d’une révision.

Tableau 1. Vue d’ensemble des opinions du personnel enseignant au sujet des PEI

Tableau 2: Plans d’enseignement individualisés
Afin de déterminer s’il y a des différences significatives entre les opinions du personnel enseignant par rapport aux cinq variables indépendantes, c’est-à-dire l’âge, le sexe, le niveau d’enseignement, le nombre d’années d’expérience et le nombre de cours suivis, nous avons procédé à une analyse de variance (ANOVA). Les douze énoncés du questionnaire ont été soumis à cette analyse qui démontrent quelques différences significatives.

**Tableau 3. Position du personnel-enseignant à l’égard des différents aspects du PEI**

<table>
<thead>
<tr>
<th></th>
<th>Utilise le PEI</th>
<th>Consulte le PEI</th>
<th>Attentes sont pertinentes</th>
<th>Inf. reflète le besoin des élèves</th>
<th>Outil péd. utile</th>
<th>Diagnostic est précis</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>11 (3,4%)</td>
<td>11(3,4%)</td>
<td>11; 3,4%</td>
<td>8; 2,5%</td>
<td>12; 3,7%</td>
<td>8; 2,5%</td>
</tr>
<tr>
<td>2</td>
<td>26 (8,1%)</td>
<td>34 (10,6%)</td>
<td>27; 8,4%</td>
<td>21; 6,5%</td>
<td>29; 8,9%</td>
<td>23; 7,1%</td>
</tr>
<tr>
<td>3</td>
<td>62 (19,3%)</td>
<td>43 (13,4%)</td>
<td>60; 18,6%</td>
<td>55; 17,0%</td>
<td>76; 23,2%</td>
<td>59; 18,2%</td>
</tr>
<tr>
<td>4</td>
<td>97; 30,1%</td>
<td>90 (28,0%)</td>
<td>94; 29,1%</td>
<td>98; 30,2%</td>
<td>106; 32,4%</td>
<td>139; 42,8%</td>
</tr>
<tr>
<td></td>
<td>126; (39,1%)</td>
<td>144; (44,7%)</td>
<td>131; (40,6%)</td>
<td>142; (43,8%)</td>
<td>104; (31,8%)</td>
<td>96 (29,5%)</td>
</tr>
<tr>
<td>total</td>
<td>322; 100%</td>
<td>322; 100%</td>
<td>322; 100%</td>
<td>324; 100%</td>
<td>327; (100%)</td>
<td>325; (100%)</td>
</tr>
</tbody>
</table>

**Tableau 4. Pertinence du PEI**

<table>
<thead>
<tr>
<th></th>
<th>Stratégies sont pratiques</th>
<th>Connaissances des anomalies sont adéquates</th>
<th>Revoir les normes du PEI</th>
<th>PEI est un outil péd.</th>
<th>PEI est une tâche administrative</th>
<th>Éducation répond aux besoins des élèves</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>11 (3,4%)</td>
<td>8 (2,6%)</td>
<td>8 (2,6%)</td>
<td>16 (4,9%)</td>
<td>23 (7,1%)</td>
<td>74 (22,9%)</td>
</tr>
<tr>
<td>2</td>
<td>33 (10,3%)</td>
<td>29 (8,9%)</td>
<td>21 (6,8%)</td>
<td>32 (9,9%)</td>
<td>40 (12,4%)</td>
<td>99 (30,7%)</td>
</tr>
<tr>
<td>3</td>
<td>82 (25,5%)</td>
<td>82 (25,2%)</td>
<td>82 (26,5%)</td>
<td>93 (28,7)</td>
<td>80 (24,8%)</td>
<td>96 (29,7%)</td>
</tr>
<tr>
<td>4</td>
<td>113 (35,2%)</td>
<td>136 (41,8%)</td>
<td>85 (27,4%)</td>
<td>91 (26,3%)</td>
<td>77 (23,8%)</td>
<td>40 (12,4%)</td>
</tr>
<tr>
<td>5</td>
<td>82 (25,5%)</td>
<td>70 (21,5%)</td>
<td>114 (36,8%)</td>
<td>92 (28,4%)</td>
<td>103 (31,9%)</td>
<td>14 (4,3%)</td>
</tr>
<tr>
<td>total</td>
<td>321 (100%)</td>
<td>325 (100%)</td>
<td>310 (100%)</td>
<td>324 (100%)</td>
<td>323 (100%)</td>
<td>323 (100%)</td>
</tr>
</tbody>
</table>
Dans le premier énoncé du questionnaire concernant l’utilisation du PEI, l’analyse de variance ne démontre qu’une seule différence entre les opinions des enseignants et enseignantes. Cette différence s’explique par le niveau d’enseignement. Le personnel enseignant du niveau primaire-moyen utilise plus fréquemment le PEI et cette différence est significative [F= (2,293) = 3,724; p > 0,05]. Aucune différence significative n’a été observée au sujet de la consultation du PEI et de la pertinence des attentes indiquées. Par contre, l’énoncé « l’information dans les PEI reflète les besoins des élèves » s’est avérée importante pour tous les répondants [x = 4,06], et une différence significative a été observée entre les groupes [F= (2; 271) =4,440; p> 0,05]. Ce sont les répondants du cycle primaire-moyen qui expliquent cette différence et ce sont eux qui ont suivi le plus de cours en enfance en difficulté: 72 au total comparativement à 28 pour ceux et celles du moyen-intermédiaire et 40 pour ceux de l’intermédiaire supérieur.

L’ensemble du personnel n’a pas une opinion très forte au sujet de l’utilité du PEI, [x = 3,80], mais les différences d’opinions sont tout de même significatives [F= (2, 312) = 4,562; p > 0,05]. Cette variance est liée aux répondants du cycle primaire-moyen qui croient davantage à l’utilité du PEI. L’analyse de variance ne dénote aucune différence importante entre les répondants en regard à l’énoncé traitant de la précision du diagnostic de l’élève telle indiquée dans le PEI. Les stratégies d’enseignement proposées au personnel enseignant sont jugés plus ou moins favorablement [x= 3,69], et seul le cycle explique, une fois de plus, les différences significatives observées [F= (2; 267) = 3,282; p > 0,05].

Les connaissances sur les anomalies des enfants en difficulté sont liées directement au nombre de cours suivis par les répondants et les différences sont significatives [F = (3, 212) = 4,730; p> 0,05]. Il faut remarquer qu’il y a ici deux fois plus de répondants du cycle primaire-moyen qui détiennent la partie 3 de la qualification additionnelle du ministère de l’Éducation de l’Ontario. Plus de 63% des répondants croient que le MÉFO devrait réviser les normes du plan d’enseignement individualisé, mais l’analyse de variance ne démontre aucune différence significative par rapport au niveau
d’enseignement.

Les répondants devaient aussi donner leur opinion sur le PEI, à savoir s’ils le voyaient comme un outil pédagogique et/ ou une tâche administrative. En examinant les données, nous pouvons conclure qu’il n’y a pas de lien entre l’attitude de ceux-ci face à ces deux énoncés. Tous les « Significatifs F » sont supérieurs au seuil de 0,05. Les moyennes (Voir tableaux 1 et 2) ne montrent pas de positions fortes sur l’échelle de 1 à 5. Quant au dernier énoncé au sujet du système d’éducation en Ontario, la valeur du F (3,129; > 0,05) indique qu’il y a un lien significatif entre les répondants et le cycle d’enseignement. Les répondants du cycle primaire-moyen pensent que le système d’éducation répond favorablement à l’ensemble des besoins des élèves même si la grande majorité des autres répondants des autres cycles (Voir tableau 4) ne sont pas d’accord avec cet énoncé.

Discussion

Les résultats suggèrent que sur les douze énoncés, la moitié seulement est justifiée par les variables justificatives proposées. Nous constatons que deux des variables indépendantes, soit le sexe et le nombre d’années d’expérience, s’avèrent peu pertinentes en rapport avec l’opinion des répondants. En effet, les analyses démontrent un lien insuffisamment fort (p > 0,05) et, par conséquent, elles n’expliquent aucunement de façon significative les différences d’opinions exprimées par ceux et celles qui ont participé à la recherche. Par contre, la variable relative au cycle d’enseignement est celle qui émerge comme l’explication principale des six énoncés qui ont été jugés significativement favorables. Elle n’est dédoublée qu’une seule fois, dans l’énoncé relatif à la connaissance des anomalies, par le facteur portant sur la formation additionnelle en enfance en difficulté. Pour les propositions qui n’ont reçu aucune justification significative, il faudra chercher les explications ailleurs. L’une des explications possible réside dans la composition de l’échantillon. En effet, le personnel du cycle primaire-moyen est le plus représenté dans le groupe. Parmi tous les répondants, 47,5% enseignent au niveau primaire-moyen, 21, 9% sont au niveau moyen-intermédiaire et 26,1%
Enfance en Difficulté

sont dans les classes intermédiaire-supérieur. Les résultats obtenus traduiraient donc, dans une large mesure, l’opinion des enseignants du primaire-moyen; il n’en demeure pas moins cependant que ce sont eux qui s’inspirent davantage des PEI dans leur enseignement. Cette constatation n’est pas négligeable car l’intervention pédagogique est fortement encouragée dès que les difficultés sont dépistées au primaire-moyen puisque les chances de réussite de l’intervention sont plus élevées quand les enfants sont jeunes et qu’ils sont plus réceptifs aux modifications proposées dans leur enseignement. Même si un lien significatif n’a pas été démontré que le PEI est considéré d’abord et avant tout comme une tâche administrative, la recherche laisse penser qu’elle va dans le sens des études de Bateman (1995) qui souligne que le PEI est vu par certains enseignants comme une charge de type bureaucratique. Aussi, en pensant à ce que les résultats montrent ici à propos des normes du PEI, par exemple, nous pouvons dire qu’ils confirment également les travaux de Thomson et Rowan (1995), lesquels attirent l’attention sur l’appréhension que le personnel enseignant ressent au sujet de la clarté des contenus du PEI.

Sur le plan méthodologique, il importe de souligner que le personnel enseignant n’a été consulté qu’avec un seul outil d’enquête: le questionnaire. Les données issues du groupe focus n’ont pas été incorporées dans l’étude, mais ont plutôt servi à valider le questionnaire. Celui-ci comportait une partie « commentaires » qui a été peu exploitée par les répondants. Le questionnaire a connu un succès énorme et a permis de récolter des données empiriques qui ne sont pas toujours accessibles dans les études actuelles sur les PEI. Toutefois, bien que nous reconnaissions que l’usage d’un seul questionnaire suffit à comprendre de façon approfondie les représentations des enseignants, il suscite néanmoins des pistes intéressantes. Une triangulation des méthodes d’enquête serait nécessaire pour poursuivre les pistes ouvertes par cette recherche.

Enfin les résultats obtenus confirment l’intérêt de la théorie de la représentation sociale utilisée dans cette étude. L’hypothèse posait la prémisses qu’il existait un lien plus marqué entre les représentations que font les enseignants du PEI d’une part, et leur expérience scolaire et leur contexte de pratique, d’autre part. Nous mettons ainsi un doute sur l’influence déterminante des facteurs personnels comme l’âge ou

**Conclusion**

Les études actuelles sur les PEI ont tendance à projeter les désirs aussi compréhensibles soient-ils des auteurs à propos d’un PEI idéal, mais elles laissent dans l’ombre les représentations des enseignants qui sont pourtant un facteur critique dans la réussite de l’éducation des enfants en difficulté. La présente recherche a voulu souligner l’intérêt de comprendre cette réalité. En étudiant l’opinion du personnel enseignant à propos des PEI, nous avons estimé que cette dernière s’exprime en lien avec l’ancrage des enseignants dans un contexte de pratique, qui, d’une manière ou d’une autre, colore leur point de vue. Le recours à l’enquête par questionnaire a permis de donner une assise empirique à cette manière d’aborder le regard des enseignants et enseignantes franco-ontariens. Les résultats obtenus sont éclairants et pourraient avantageusement servir de point de départ à des études utilisant des méthodes de recherche plus poussées, ces dernières pouvant ou non faire appel à une variété de méthodes. Les indications obtenues à propos du personnel enseignant du primaire-moyen nécessiteraient d’être mieux documentées en les replaçant, par exemple, dans le contexte particulier de ce cycle scolaire et dans celui, plus général, de la réforme scolaire en Ontario.

**Références**


Rapport d’évaluation. ERIC: ED 397580.

Revue française de pédagogie. Situations de handicaps et institution scolaire. No. 121 (numéro thématique).


Media Review

*Living planet report, 2006*
by the World Wildlife Fund


*Stern review: The economics of climate change*  
(Executive Summary).  
by Nicholas Stern


Reviewed by Michael O'Sullivan, Brock University

---

Educating for a Sustainable Future: Indispensable Resources

It is hard not to become discouraged whenever an environmental “state of the planet” report is issued. Inevitably, we seem to be that much closer to an ill-defined but much-discussed “ecological tipping point” past which we are told there is no return. The obstacles to global ecological sustainability are particularly overwhelming in light of the triple ecological challenge: the attachment of consumers, especially in the developed countries, to their unsustainable and insatiable lifestyle; the relentless pursuit of corporate profit by companies that serve the consumer market; and the legitimate desire of the socially and economically marginalized populations of the developing countries for an equitable share of the world’s resources.

For educators this apparently bleak state of affairs presents a
problem of how to create in their students a globally-oriented awareness that confronts this inequitable and unsustainable situation in a way that
  · avoids a doom-and-gloom scenario of the inevitability of the “tipping point”;
  · addresses the legitimate claim that the residents of the developing world have to decent standards of living, and
  · does not reduce the issue to approaches that, while worthwhile in and of themselves, have the effect of limiting student awareness of and actions around this issue to their responsibility to practice “the three Rs.”

Teaching *education for a sustainable future* requires educators to be ecologically literate. That, in turn, requires educators to be able not only to articulate the problems facing our planet but also to have a clear understanding of potential solutions that they can confidently share with their students.

Recently, two reports were released that together provide current data on the ecological challenges that we face and which also show that there are possible solutions to these issues. In mid-October the World Wildlife Foundation (WWF) issued its *Living Planet Report, 2006*. Days later the Government of the United Kingdom issued the *Stern Review: The Economics of Climate Change*, written by a team chaired by Sir Nicholas Stern, the former chief economist of the World Bank. Both reports deserve our attention, although the latter, at almost 600 pages is best read in the form of its 27-page Executive Summary.

The WWF report is written in language that most readers will feel comfortable with while the Stern report, although by no means inaccessible, is written in the language of economics. It is, however, precisely its economic orientation that makes Stern’s *Economics of Climate Change* an important and compelling read for educators who seek to guide their students towards the realization that there are solutions to the state of the planet’s ecology and that they, both as consumers and as active global citizens, have a role to play in bringing about those solutions.

The WWF *Living Planet Report, 2006* uses the metaphor of the ecological footprint as a starting point. The authors remind us that
if it is equitably shared the biocapacity of the world provides each inhabitant of this planet with 1.8 hectares to satisfy our needs. At that level of resource use, the earth can replenish what we take. If our individual consumption requires more of the earth’s capacity than 1.8 hectares, then we are contributing to consuming more resources than the earth can supply. On average the earth’s inhabitants are consuming in excess of that – 2.2 hectares – which puts the globe in a deficit situation. The average ecological footprint in some countries is well below 2.2 hectares, indeed well below 1.8, while others such as the four worst offenders (the United Arab Emirates, the U.S., Finland and Canada, in that order) consume in excess of 7 hectares per person. It should also be noted that the 1.8 hectares per capita is not a fixed number. There are only so many hectares of productive land and sea available on the planet. As the global population grows there is less productive area per capita available. Furthermore, if through global warming and other ecologically destructive phenomenon once productive areas become unproductive (farm land because of drought, for example, or the depletion of fish stocks), the amount of productive area is reduced.

The WWF Report notes that if we project present patterns into the future (what both the WWF and the Stern Review refer to as the business as usual model) the deficit between the regenerative capacity of the planet and what resources we use will only get worse. This is the case because not only will the world’s population increase (thus reducing the number of hectares per capita available) but some of the countries in the developing world that now consume less than 1.8 hectares per capita (e.g., China and India) are following development models that will increase their footprint well over the 1.8 hectare threshold in ways that are ecologically destructive (e.g., the massive emission of greenhouse gases that will contribute to further global warming).

Fortunately, this doomsday scenario is not inevitable. The WWF Report argues that there are two alternatives to the unsustainable “business as usual” scenario. The first of these they call the slow shift to a sustainable balance which would, by 2100, bring the world average per capita ecological footprint down to 1.5 hectares from the present 2.2. This, of course, is still an average and may well
include wide inequities among countries as is presently the case. The second strategy is what the Report terms rapid reduction. This would accelerate the achievements of the slow shift approach and significant improvement toward sustainability would be achieved much earlier, by 2050, and a 40% decrease in humanity’s biological footprint would be achieved by 2100.

The problem with these scenarios is that the WWF offers no strategy to achieve them. We can all agree that the globe cannot sustain the pattern of growth we are presently experiencing but given the factors that drive our models of unsustainable development, what are the incentives to achieve both ecological sustainability and a more equitable fulfillment of basic human needs?

An answer to this complex question is suggested in the Stern Review. The WWF Report takes the ecological footprint as its starting point. The Stern Review takes climate change caused by our ever-increasing consumption of fossil fuels as its starting point. These two concepts – the ecological footprint and climate change – are closely linked.

The Stern Review makes the very pragmatic argument that the economic costs of doing nothing and ignoring the human-made causes of global warming are greater than investing in measures that would reduce the greenhouse gas emissions caused by burning fossil fuels. Said another way, and using language that corporations and the governments that support them understand, Stern argues that, “The evidence shows that ignoring climate change will eventually damage economic growth” (p. ii). As compelling as arguments about the plight of the polar bear on diminishing ice floes may be for environmental activists, for corporate and government leaders a more convincing argument is that one of the consequences of global warming will be a 20% drop in global consumption (Stern, p. x), something that would be catastrophic both in human terms (Stern notes that the worst hit would be the vulnerable less developed countries [p. vii]) and from the point of view of the corporate bottom line. How is the consumption of consumer goods, including basic food, shelter, and clothing that sustains the poor, related to global warming and why should corporations see it as being in their interest to change those practices that contribute to it?

Stern argues, as convincingly as any environmental activist
would, that global warming would have devastating effects. He notes the negative effects on the oceans, on rainfall, and on temperatures:

- With respect to the oceans (p. vi), not only does global warming threaten fish stocks and the human economic activities related to it, but it will lead to a rising of ocean levels which will result in the partial or total destruction of coastal cities and the loss of agricultural land which would result in mass starvation and the dislocation of millions of coastal residents. The fate of New Orleans and its residents is fresh in our memories. We can only imagine what would have happened if that city had been located in Bangladesh, which would have no resources to cope with the aftermath.

- Changes in rainfall (p. vii) would affect available supplies of fresh water for human consumption and for agriculture. This, too, would result in dislocation and starvation, particularly in the already poor and vulnerable countries of the global South and also for the wheat belt in North America.

- Temperatures in Europe (p. viii), which has suffered extraordinary heat waves in recent years, will continue to rise with global warming. The impact on present tropical climates should temperatures rise even a few degrees would be catastrophic.

These are only three of the many impacts noted in The Stern Review. Under such circumstances it is little wonder that Stern and his team of economists predict that incomes would fall and, especially in the global south, consumption would decline. It is not unrealistic to predict that political unrest would erupt, leading to many humanitarian disasters like the one in Rwanda a decade ago or in Darfur today. In addition, there would be an escalation of militarization, as nations with fresh water or productive agricultural land would feel threatened by their desperate neighbours.

While the WWF presents concerns such as these in the framework of a humanitarian call to action, albeit a self-interested one as we all share in the fate of the planet, The Stern Review presents concerns in terms that speak to corporate and political self-interest. Furthermore, the solutions Stern proposes are made in terms that corporations and politicians understand: investment in new technologies and production techniques, risk reduction, political (e.g., market) stability; and potential for long run growth. Given the fact that
governments, including Canada, China, Russia, and the U.S.A., feel they can ignore Kyoto with impunity, clearly the environmental movement does not have the clout needed to insure that even democratic nations are doing the minimum required to reverse the rush to ecological disaster. Sometimes breakthroughs come from unexpected quarters. Faced with the predictable loss of their markets and with solutions including such tools as carbon taxes to punish offenders and reward innovation, the Stern recommendations, couched in business terms but oriented to ecological sanity, may be a huge step in the right direction.

And how does the educator link these apparently corporate- and government-driven decisions about investments and alternative energy sources into potential civic involvement by student-citizens? Corporations will not do these things on their own nor will the people nor the governments in developing countries unless they are given guarantees that ecologically sound policies will allow for growth and a rise in their standard of living. As Stern puts it “an effective response to climate change will depend on creating the conditions for international collective action” (p. xxii). Competition requires a level playing field. In today’s world that level playing field has to be global. Only when all concerned, whether they be governments or private transnational corporations, are simultaneously encouraged by incentives and obliged by national and international law to behave in certain ways will they make any significant move that contributes towards reversing present destructive ecological trends and take steps to reduce the ecological footprint of the residents of the developed countries while creating the conditions to meet the legitimate material needs of the world’s poor.

Stern, as a former senior official of the World Bank, knows full well the complexities of world politics and the power relations upon which they rest. When he calls for urgent and concerted international action, he is fully aware of the political and corporate resistances. He is also knows what can galvanize these stakeholders into action. Corporations require a guarantee that new ecologically sound initiatives must be undertaken by all because the first to do so will be at a competitive disadvantage due to higher costs. This is where governments and international agencies such as the United Nations are
important. Stern provides a strategy which meets human needs, protects corporations’ need to survive, and addresses the state of the world’s environment. An informed citizenry in the countries that comprise the top 30 worst offenders to our global environment as measured by the size of their ecological footprint (WWF, p. 14) has a key role to play in convincing their governments to begin to create the conditions whereby these steps to ecological survival are taken. Educators can be important in making this happen.