A NEW LOOK AT PUBLIC ASSURANCE: IMAGINING THE POSSIBILITIES FOR ALBERTA STUDENTS



everyday assessment tools to support student learning This discussion paper and accompanying classroom samples have been funded in part by a research grant from the Alberta Teachers' Association.

More information available at www.aac.ab.ca (Advocacy and Research tab)

© 2012

CONTENTS

- 2 Foreword
- 3 Preamble
- 5 Exploring the Role of Accountability
- 6 Time for a Change: Limitations of Existing Large Scale Assessment Practices
- 8 A New Model of Public Assurance
- 15 Proposed Implementation Plan and Next Steps
- 17 Appendix A: Review of Large-Scale Testing Reforms in Various International Jurisdictions
- 19 Appendix B: Recommendations for Provincial Assessments
- 24 Appendix C: Essential Conditions to Support Implementation
- 26 References
- 28 Bibliography





FOREWORD

In 1988, a quarter of a century ago, the first meta-analysis of global research on educational assessment was published. This study of the links between assessment, learning and achievement revealed for the first time what many people interested in this field – myself included – had always suspected. To use an agricultural analogy, 'Weighing the pig doesn't make it grow.'

Since then the evidence about what kinds of educational assessment have the most positive impact on educational outcomes has kept coming, inexorably. Whole cohort testing, once the norm in many jurisdictions, has been questioned, adapted, reviewed, and questioned again. The issue is not whether we need information about the learning and achievement of our children and young people, but what kind of information we need, and how best to gather it. The first requirement of world-class public assurance and accountability is world-class information about student achievement.

On the large scale of a province, or the small scale of a school, we must pursue valid information about student learning. How accurate is the information relative to the kind of learning we aspire to achieve? Secondly, and most importantly when assessment information may be used to compare outcomes from one place to another, one group to another, or a cohort of students compared with the same cohort previously, we need to reduce the variables that infuse educational assessment. Thirdly, we must pay attention to the cost and logistical implications of information gathering. Is the benefit to be derived from the assessment process worth the cost, in terms of both dollars and time? To use another North American phrase, 'How do we get more bang for our buck'?

This very timely paper reviews the decisions facing Alberta, and how other states and jurisdictions are currently finding their way through the maze. It identifies the skills and capacities our students need to be successful learners and earners in the twenty-first century, different in many respects from the skills of previous decades. Many of these skills do not lend themselves to 'data capture' through traditional forms of testing: they are manifest not in answers to a pencil and paper test but in student responses to on-going classroom challenges, observed by their teachers. Time and money spent on using inappropriate methods of data capture could be time and money wasted.

To satisfy public assurance without relying on outdated testing procedures, we will need to look carefully at the variables involved and find new ways of reducing their impact, using proven strategies from elsewhere around the world. The two greatest variables are the design of the assessment task itself, and the interpretation of assessment outcomes. The design of tasks can be supported by central professional expertise, and through specific training for teachers in this aspect of their work. The problem of variable interpretation of results is tackled by requiring teachers to share their expectations and evidence of learning, and then to 'moderate' their judgements, reaching a collective rather than an individual decision.

Information about student achievement provides the building blocks for the structure of public assurance. If the information is of poor quality, the structure will always be weak. Once the quality of information is assured, then desirable and effective structures can be built.

These aspirations are worthy in themselves, but the proposals in this paper go further. Learning from others' experience, Alberta can develop a system of public assurance that not only measures the impact of schooling but actually strengthens both classroom learning and the schools and local organizations that provide it.

Involving teachers collectively in assessment and evaluation provides highly effective professional development, as teachers learn from each other and share successful practice. Involving schools and their communities in the process of school and district review and evaluation serves the same purpose, sharing the most successful and effective strategies for improving educational outcomes for all our students.

Research and experience over the last twenty-five years has shown the way forward. Now we need to help our communities and our tax-payers understand these irresistible messages. Involving teachers, parents, students and the community in the process of assessment and evaluation is the key to a form of public assurance that serves to improve and develop our students and our schools, not just to measure them.

Ruth Sutton, May 2012

PREAMBLE

Although the conversation about educational accountability in Alberta is not new, recent events have sparked a renewed interest in the topic.

- In March of 2009 the Alberta Legislature passed Motion 503, recommending the elimination of the Grade 3 Provincial Achievement Tests (PATs) and giving consideration for alternative methods to assess student achievement at the grade three level. Although the motion was passed, it was a private member's motion and as such, the government is not obligated to act on it.
- In the spring of 2011, Alison Redford, while campaigning for the leadership of the Conservative party and ultimately the office of Premier stated her position in support of finding an alternative to the grade three tests.
- In January of 2012, Education Minister Thomas Lukaszuk in his 10 Point Plan for Education committed to "reviewing provincial achievement tests, first so that we can find better ways to understand how students, schools and the system are doing and then so we can use that knowledge to improve students' success" (Alberta Education, 2012).
- Alison Redford, newly elected Premier of Alberta responded to a question in regard to eliminating Grade 3 and Grade 6 Provincial Achievement Tests by stating that "...we're going to transition out of those. We want to make sure we have a system in place so that we can track progress in students, so that if there are students that are falling behind, we can identify that, but I don't think Provincial Achievement Tests are the way forward for us, and we've asked Education to look at what alternative models we could put in place" (CBC Alberta at Noon, May 10, 2012).
- Jeff Johnson, Alberta's new Education Minister stated that "...as we slowly transform the educational system, that has a lot of implications for how you assess ... The PATs, in their current form, will not be what we are using 10 years from now, is my prediction, but there's still an important requirement to have some form of standardized assessment across the province" (Edmonton Journal, A4, May 15, 2012).

Three years have elapsed since Motion 503 was passed. Although numerous questions have been raised in the Legislature in regard to when some action would be taken in response to the motion, former Education Minister David Hancock frequently stated that an alternative would have to be found before the tests would be eliminated. Alberta's new Premier and Education Minister both indicate that the need for some level of public assurance remains, while acknowledging that the current model is insufficient. The vision of what that new framework might look like has yet to be articulated.

Alberta has much to gain by ensuring that our young learners acquire the requisite knowledge, skills and attitudes to ensure a solid foundation for future learning. Yet how can we know that students are 'ready' for grade four? Is it possible to design alternative assessments that can both support learning and at the same time, assure the public that Alberta students are receiving a high quality, world class education? It is the view of the Alberta Assessment Consortium (AAC) that it is not only possible, but highly desirable. Assessment authors and researchers from around the world agree.

Is it possible to design alternative assessments that can both support learning and at the same time, assure the public that Alberta students are receiving a high quality, world class education?

Founded in 1994 as a not-for-profit organization, the AAC is dedicated to the enhancement of quality classroom assessment practices. AAC occupies a unique position among the education partners in Alberta as it is not funded by government but rather by membership. As such, AAC is able to offer an independent voice, providing informed responses to assessment topics that impact student learning. AAC contributes to assessment literacy in Alberta by developing a broad range of classroom assessment resources and professional learning materials; facilitating lateral networks among school jurisdictions and education partners; and engaging in action research and inquiry.

This discussion paper is offered as an invitation to explore possibilities for a new look at public assurance through a class-room based diagnostic assessment approach to replace the current grade three Provincial Achievement Tests. It is not intended to be prescriptive nor does it suggest that the views expressed herein represent the views of individual members of the consortium. The intent of this paper is to promote public dialogue in regard to this important topic.

The paper begins with an exploration of the notion of accountability, and then considers the limitations of existing large scale assessment practices. Credible alternative assessment measures are presented, along with a proposed implementation sequence to ensure success. Samples of these classroom based assessment measures in action will be available on the AAC website at www.aac.ab.ca under the Advocacy and Research tab.

As we embark on the transformation of education in Alberta, it stands to reason that changes to curriculum and pedagogy will require corresponding changes to assessment practices. In the spirit of inquiry, we invite you to explore with us some new possibilities for assuring the public that our students are receiving the very best education possible. New ideas and information will surely prompt further questions and subsequent explorations. We invite you to join the inquiry by imagining the possibilities for Alberta students.

EXPLORING THE ROLE OF ACCOUNTABILITY

All citizens benefit from a quality public education system. As expenditures for education constitute a significant portion of the provincial budget, the public has a right to know that this investment is producing optimal results. To this end, large scale assessment systems have been developed for the purpose of assuring the public as to the quality of the education system in Alberta. Notwithstanding the intent, accountability measures are always influenced by the biases, perceptions and decisions made by those designing the system. The quest is not to be in search of perfect accountability, if indeed that state exists, but rather to seek intelligent accountability (Sahlberg, 2011). Thus, any educational accountability system should ensure that what is measured is what matters most, and that the results guide the system towards continuous improvement in support of student learning, not just student performance on a narrow range of measures.

The Ministry has sought to broaden the measures within Alberta's Accountability Pillar by including measures of student and parent perceptions within the affective domain such as safe and caring schools; preparing for lifelong learning, employment and citizenship; and parental involvement in addition to measures of student achievement. However, data within the affective domain gathered through standardized means may not be sufficient proxy for authentic community voice. In the recent Alberta Education *Inspiring Education* consultations, a desire was expressed for less standardization and more community involvement.

In 2030, if Alberta is to truly foster learning excellence, there will be no "one-size-fits-all" approach. What is taught, how it is taught, and how the community is engaged will reflect what is valuable to the community. Similarly, accountability processes will reflect the appropriate degree of complexity and formality required by teachers, governors, leaders, managers and funders (Alberta Education, 2010, p. 37).

In the search for intelligent accountability, we cannot ignore the value of community engagement, student engagement, or other so called 'soft' indicators. As important as they are in their own right, they also serve as predictive indicators of student success (Murgatroyd & Sahlberg, 2010) and, as such, merit our considered attention.

While not suggesting an absence of accountability, leaders in the field of educational reform suggest repositioning or reframing traditional notions of accountability. Fullan states that "...no system in the world has ever achieved whole system reform by leading with accountability" (2011, p. 19). Rather, leading with capacity building and trust will result in "greater lateral accountability among peers, which is absolutely critical for whole system reform." Hargreaves (2009) suggests we

...put responsibility and transparency first and use accountability as a backup to check, in sampling terms, whether we truly achieve the quality we claim" (p. xvii in *Educational Accountability: Professional Voices From the Field*, K. Gariepy, B.L. Spencer, & J-C. Couture Eds.).

Thus, the desired results of public assurance and improved student performance can be achieved in ways other than through a focus on external accountability.

...the desired results of public assurance and improved student performance can be achieved in ways other than through a focus on external accountability.

TIME FOR A CHANGE: LIMITATIONS OF EXISTING LARGE SCALE ASSESSMENT PRACTICES

By their design, large scale measures such as the current Alberta Provincial Achievement Tests (PATs) are only able to measure a limited number of educational outcomes. As Alberta moves towards a competency focused curriculum, the current provincial testing model will be able to provide less and less information about aspects of curriculum that are considered to be essential. These essential elements include critical thinking, problem solving and decision making; creativity and innovation; social, cultural, global and environmental responsibility; communication; digital and technological fluency; lifelong learning, personal management and well-being; and collaboration and leadership (Alberta Education, 2011).

Numerous researchers have brought to light the negative impact of testing on student motivation and learning (Assessment Reform Group, 2002; Collins et al., 2010; Doecke et al., 2010; Hamilton, 2010-2011; Jones & Egley, 2004; Volante, 2004; Wiliam, 2010). They lead us to ask questions that challenge educators to rethink the value of large scale testing as the primary means of accountability.

DOES TESTING RAISE THE LEVEL OF ACHIEVEMENT?

- Researchers have "ample evidence that teaching to the test is not the best way to improve test scores" (Bew, 2011)
 but it is very difficult for teachers to resist doing just this. Pressure on teachers to have students do well on the test may
 prompt them to fixate on test format and test preparation, furthering a misbelief that practicing a specific format will
 improve test scores. These strategies are not conducive to the development of educated individuals.
- Large scale test results typically come too late to be of any value in meeting the learning needs of individual students, and teachers outwardly acknowledge that some students do not do well on tests due to anxiety issues. Therefore, such tests do not provide an accurate picture of the child, but rather only a brief snapshot of a limited and specific contextual moment in time.

ARE STUDENTS MOTIVATED BY TESTS?

- There is strong evidence of the negative impact of testing on student motivation, particularly on students who achieve low marks. Only students who are confident of their success on tests become more persistent and develop a more positive self-perception within a testing situation.
- By contrast, low achievers become overwhelmed and discouraged, an impact that may have significant consequences for their futures. Therefore the gap between low achievers and high achievers potentially widens.

DO LARGE SCALE TESTS LEAD TEACHERS TO ADDRESS THE BREADTH OF ACADEMIC LEARNING OUTCOMES?

- The design of current large scale assessments excludes many of the most valued goals of schooling for the future and instead, focuses on concepts that can be easily measured and machine scored.
- Repeated practice tests lead students to believe this is what is important and may encourage a focus on test taking strategies over persistence towards mastering higher order thinking skills.
- If there is extensive teaching to the test, the reliability of any single test as an indicator of attainment of standards must be questioned.

The current system of large scale assessment in Alberta that leads to the practice of rating and ranking schools implies a cause/effect relationship between quality of the educational experience and school rank. It fails to take into account the

- appropriateness of the assessment in relation to the developmental needs, particularly of young children;
- individual student learning rates and styles;
- impact of socioeconomic factors;

As Alberta moves towards a competency focused curriculum, the current provincial testing model will be able to provide less and less information about aspects of curriculum that are considered to be essential.

- growing linguistic and cultural diversity within Alberta, including the proportion of English Language Learners within a school population;
- distorted results that can ensue within small populations in rural jurisdictions;
- · world-wide research into curricular design with its focus on 21st century skills and competencies; and
- priorities resulting from contextual variables within communities.

The manner in which results from large scale assessments are shared has potential consequences for decisions that impact jurisdictions, schools, teachers, and possibly even students. It is critical that information shared with the public ensures clarity about the purpose for large scale assessments, what can validly be measured, and what can only be inferred from the results.

These [large scale] assessments are designed to determine what students know, can do and can articulate in relation to what is to be learned. They do not collect enough information to give a valid and complete picture of everything students know and are able to do at any point in time, in relation to all the standards or learning outcomes they are to learn. Large-scale assessments can only provide a snapshot of some of the learning. They are better designed to describe what groups of students are able to do.

Classroom assessment is quite different. Teachers and students collect a large amount of evidence over time from multiple sources. It is designed to account for all that is to be learned by individuals – student by student. When done well, classroom assessment is better able to give a more valid and reliable accounting of a student's learning (Davies, 2011, p. 51).

In its original design within a sampling approach, provincial testing had the potential to provide feedback to the system as a whole. However, over the years Alberta Provincial Achievement Tests have become universal in their administration, providing a profile of individual student achievement at the targeted grade levels. Access to provincial testing data by third party organizations has the potential to lead to misuse of data, thereby undermining the education system within Alberta school jurisdictions.

As Alberta moves towards a competency focused curricular framework, it is essential that assessment and accountability measures are consistent with this direction. Findings from the *Inspiring Education* dialogue highlight that "...new processes are required to continuously evaluate whether the education system is "living" the vision of Albertans" (Alberta Education, 2010, p. 37).

A number of factors suggest that both the will and the expertise now exist to move from the current model of accountability to a model focused on public assurance.

- Alberta educators have been exposed to an extensive worldwide research base in the field of classroom assessment, including both formative and summative assessment practices. This has been evidenced by the number of Alberta Initiative for School Improvement (AISI) projects over several cycles that have focused on classroom assessment practices.
- The Alberta Assessment Consortium continues to expand its reach within the province and beyond, creating robust classroom assessment tasks based on Alberta programs of study and providing quality professional learning opportunities and resources to support the building of classroom assessment capacity.
- Alberta Education has demonstrated a commitment to the important role professional learning plays during implementation of new curricula and other initiatives.
- The Premier and Minister of Education have, as recently as May 2012, indicated the need to explore changes to the provincial achievement testing program (CBC Alberta at Noon; Edmonton Journal).

These factors contribute to an educational system that is well equipped to explore a new public assurance model of accountability.

A NEW MODEL OF PUBLIC ASSURANCE

The model being proposed in this paper does not advocate less accountability, but rather, a change in the way accountability is viewed and managed. It also highlights the prominent role of teachers' informed professional judgment in determining student progress by grade level which was a key principle articulated by the Building Assessment Capacity Advisory Committee, a committee convened by Alberta Education in 2008 for the purpose of considering future direction in classroom assessment within Alberta.

A number of models for effective system accountability that are also supportive of learning are currently being considered around the world, a summary of which is found in *Appendix A: Review of Large-Scale Testing Reforms in Various International Jurisdictions*. Although slightly different in their various approaches, the models adhere to similar essential components.

- An effective system should assess the most important goals in appropriate ways, not simply assess those goals that are easiest to measure. This requires the use of a variety of assessment methods.
- The system should provide specific and timely feedback to improve learning. Consistent with current research, a greater focus on assessment for learning provides teachers with information to make adjustments to instruction, making the system diagnostic in design.
- The system should be supported by a professional development component that adds to the skills and content understanding of teachers in integrating curriculum, planning and instruction for meaningful learning (McTighe & Wiggins, 2010; Wiliam, 2010).

To address these components and provide a system that will help teachers determine how well students are learning relative to provincial standards, a multiple measures approach to be administered in the fall of the grade four year is proposed. The movement to a fall administration repositions the teacher's role in the assessment process as a planner, as the assessment evidence can be used to inform instructional practice in service of improved student learning.

The proposed model would include two types of measures.

- 1. Provincial Measures of Literacy and Numeracy
- 2. Locally Determined Measures

The provincial measures of literacy and numeracy would be administered to all grade four students with the results made available for provincial sampling. Locally determined measures would also be focused at the grade four level, based on priorities determined through input from the local community. This collaborative process of determining priorities at the community level would in turn yield measures of community engagement. For the purpose of this discussion paper, the emphasis on an early fall, grade four administration provides a model that could be used to inform future consideration for redesign of provincial assessments at other grade levels.

A summary of the proposed measures follows, while *Appendix B: Recommendations for Provincial Assessments* provides a discussion of some specific examples. Samples of these measures within a classroom context can be found on the AAC website at www.aac.ab.ca under the *Advocacy and Research* tab.

1. PROVINCIAL MEASURES OF LITERACY AND NUMERACY

a. Literacy Assessments of Reading and Writing

At the beginning of grade four, teachers would administer an individual reading assessment with each student. Many teachers currently use some form of reading assessment at the beginning of the year to determine leveled reading groups. Information provided by the assessment would be used formatively to adjust instruction and assist teachers in identifying specific strategies for individuals and small groups of students.

A writing task would also be administered in the early part of the year. Topic and genre would change over time, providing a bank of student exemplars in a variety of writing styles and formats. Teachers would be able to provide

feedback on various aspects of the writing, thereby supporting student growth in the writing process on writing assessments throughout the year. The results would provide evidence as to the current performance level of writing in a specified genre, supported by criteria and exemplars.

The fall administration would provide teachers with diagnostic information to guide ongoing instruction throughout the year and the results of these assessments would be available for provincial sampling.

b. Performance Assessments in Literacy and Numeracy

Performance assessments are meaningful, real-life tasks that enable students to demonstrate what they know and can do in situations set in real world contexts. Performance tasks are well suited to assess skills such as reasoning, inquiry, issues analysis, problem solving, oral communication and digital and technological fluency. These tasks would assess the listening, speaking, viewing, and representing components of the English Language Arts Program of Studies and the reasoning, visualization, problem solving and communication process skills of the Mathematics Program of Studies that are more difficult to assess in traditional ways.

On an ongoing basis, the use of performance assessments should not be limited to language arts and mathematics, but rather encouraged as a credible assessment method to be employed throughout the year in all subject disciplines and in multi-disciplinary contexts. However, for provincial assessment purposes, it is desirable for the assessment tasks to be focused on literacy and numeracy skills rather than on specific subject related content and skills, such as might be the focus for social studies or science assessment tasks. The selected skills focus from the Language Arts and Mathematics Programs of Study for the provincial assessments would also be intentionally linked to the competencies identified in the *Framework for Student Learning* (Alberta Education, 2011).

While a broad bank of tasks in a variety of subject areas and grade levels would be available for use by teachers throughout the year, specific tasks would be developed for the fall provincial assessment administration. These tasks might be similar to those found on the Alberta Assessment Consortium website. Ideally, the development of such tasks would be a collaborative effort, combining the expertise of both the Alberta Assessment Consortium and Alberta Education Learner Assessment. The AAC brings extensive expertise in classroom based assessment practices and a grass roots connection with the field, while Learner Assessment brings experience with validation of assessment items and experience with statistical analysis of data. Collaborative teacher scoring of performance tasks could also include a professional learning component, with support from Alberta Education Learner Assessment to ensure consistency in scoring, and AAC to assist with follow-up formative assessment practices consistent with a diagnostic approach to assessment.

c. Student Reflection

Schools need to be developing learners who have the ability to address competencies for engaged thinking and ethical citizenship within a literacy and numeracy context (Alberta Education, 2011). There is evidence that having students reflect on their learning and articulate what they need to do to improve assists them to see their role in the assessment process as one of owner, rather than one of receiver of assessment information (Wiliam, 2010).

As Ruth Sutton states, "a learner is someone who knows what they know, who knows what they don't know and who knows what to do when they don't know what to do" (2005). In short, to build autonomous learners we need to transfer executive control of learning from the teacher to the students.

We must constantly remind ourselves that the ultimate purpose of evaluation is to have students become self-evaluating. If students graduate from school still dependent on others to tell them when they are adequate, good or excellent, then we have missed the whole point of what education is about (Costa & Kallack, 2003, p. 3).

Student self-reflection would not be an isolated task, but rather would be embedded within the reading and writing assessments as well as within the performance assessments. A number of means for "activating students as owners of their own learning" (Wiliam, 2011) include the use of student questionnaires; self-reflections through learning logs

or learning portfolios; and ongoing reflection as learning is taking place. Students can be asked to reflect on their abilities, understanding and attitudes about what they are reading and writing. They can discuss how and why their achievement has shown growth over time with regard to a specific piece of work, or through several pieces of work in a learning portfolio. The better students are at managing what and how they are learning, the more success they will see as learners.

Expanding the range of assessment methods will also require exploring new ways to record, capture, collect and store the evidence of learning. Digital portfolios provide options that were simply unavailable in a previous era. Technology applications can be used to capture observations of students engaged in learning; record conversations with students about their learning; display student performances and products; and share student self-reflections about their learning thus providing multiple means for informing teacher professional judgment in regard to student achievement levels.

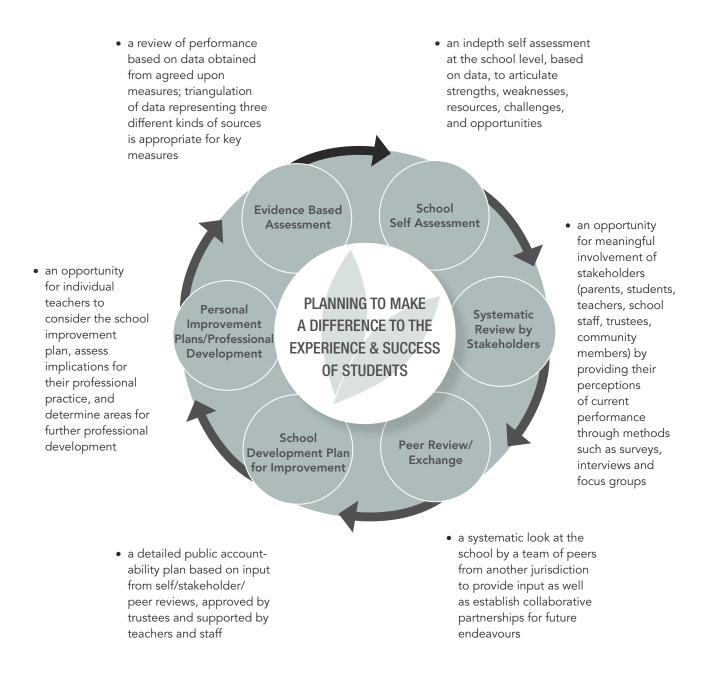
Teacher professional learning opportunities would be provided to support teachers in the administration of the measures, development of protocols surrounding the use of exemplars of student work, as well as with strategies to promote strong literacy and numeracy practices. The professional learning component has the added benefit of enhancing assessment capacity across the province.

2. LOCALLY DETERMINED MEASURES

"A school is the point of identity for students, parents and teachers" (Murgatroyd, 2011, p. 177). Students will more likely identify themselves as a student of a particular school rather than as a student of a particular school jurisdiction. As such, the school is well positioned as the focal point for public assurance purposes.

Based on an action research project launched in 2007 with schools in Livingstone Range School Division and Grande Prairie Public Schools (Murgatroyd, 2011, p. 178), a model for local school improvement planning has emerged (Figure 1: School Development Planning Cycle). This model represents a cyclical and continuous process of school improvement. Not surprisingly, the model is consistent with principles of sound assessment practice which support student learning at the classroom level. The six steps in the process are briefly described on the next page.

Figure 1 School Development Planning Cycle



Data at the conclusion of the cycle informs the beginning of the next cycle, thus seeding a continuous cycle of school improvement. This school improvement cycle relies on the focused commitment and involvement of the local community.

Within the public assurance framework at the centre of this discussion paper, locally determined measures include measures related to student learning as well as measures of community engagement. These will be explored through vignettes of two specific school communities.

a. Local Measures of Student Learning

While public perception may suggest that PAT results and other Accountability Pillar measures are standardized and therefore more reliable than local measures, a vignette from a small northern school provides insight into the power of locally determined measures of student learning.

The student population in École Providence School in Holy Family Catholic Regional Division consists of a high percentage of First Nations, Métis, and Inuit (FNMI) students as well as several students with identified special needs. The combination of low enrollment at the target grade levels for the provincial exams, a very small teacher population, and low response rates from parents on the Accountability Pillar surveys yields a distorted picture of school performance on both academic measures as well as affective measures from the Accountability Pillar. Thus, it becomes difficult for school leaders to analyze results and work towards meeting improvement targets when the inclusion or omission of a single student/teacher/parent may have a dramatic impact on results.

Accountability results aside, educators at the school were determined to help improve students' basic literacy skills. The school embarked on a school wide reading intervention program. By the end of the first year of the reading intervention, 85% of the students were reading at grade level and 50% of the students gained one or more grade levels in their reading scores. Middle school students were trained as peer tutors, assisting primary students to improve their reading. Improved attitudes towards reading were evident among the primary students as well as the middle school peer tutors (École Providence School Annual Education Plan, 2010).

Important gains in student reading levels deserve to be publically celebrated. However, this can be challenging when the public perception, as based on standardized measures is so pervasive as to render local measures of lesser import. While the statistical analysis available from the current Accountability Pillar provides a certain degree of information that can be helpful as school leaders seek to understand trends and identify areas for improvement, the usefulness of the data depends in part on the degree to which the analysis can be applied, particularly when dealing with special and/or small populations.

Many Alberta students fall outside of the regular grade level programming and as such, are either not included in the current accountability measures or if they are included, do not perform well due to their specific programming needs. Local measures of student learning provide avenues to celebrate progress and success that are not tied to grade level standards or provincial accountability measures. These may include accomplishments such as gains in student reading levels, the ability to solve problems, development of leadership skills, the ability to collaborate, and development of personal management skills.

While there is nothing currently preventing jurisdictions from reporting locally determined measures, the emphasis on the Accountability Pillar may lessen the desire of jurisdictions to invest time and resources to that end. A restructuring of the current Accountability Pillar would give new credence to local measures and celebrate the learning of all students. Changes in AISI Cycle 5, that remove the requirement to have a certain percentage of provincial measures, while embedding a focus on community engagement are positive. Moving away from provincially determined measures towards locally determined measures will require resource allocation to support leaders in selecting and determining credible measures and developing skill in collecting, analyzing and interpreting data.

b. Local Measures of Community Engagement

As the community engages in helping determine school improvement goals, authentic measures of community engagement emerge. A tremendous opportunity exists within AISI Cycle 5 to explore more deeply the impact of community engagement on student learning and school improvement.

Effective school improvement results when principals, teachers, other school staff, students, school councils, parents and other school community members work as a team to establish priorities, set goals for improvement, implement strategies to achieve those goals, and evaluate progress (AISI Cycle 5 Handbook 2012–2015, 2012a, p.3).

While support materials for AISI Cycle 5 include both a continuum and a rubric for community engagement, neither is intended to be used as a measurement tool to obtain a numerical score.

The process for determining a profile of school community engagement in a project is one of informed and collaborative judgement. Neither the rubric nor the continuum is a measurement instrument. Rather, each is a structure that helps project teams organize and analyze evidence and make judgements. This evidence is assembled from various sources, such as surveys, focus group notes, meeting minutes, and anecdotal notes (School Community Engagement Rubric and Continuum, 2012b, p. 4).

Although the requirement for community engagement is new for this AISI Cycle, Livingstone Range School Division has been working on a collaborative project aimed at the development of a process for school improvement and accountability since 2007. The voices of teachers, parents, students and the community were seen as essential in the design. A divisional framework provided a shared structure with enough latitude to honour the work of individual schools. School based researchers provided support for collection and analysis of data in order to keep the project goals on track and "ensure accountability to the division, the school, students and the community" (Adrian, 2012). The Livingstone Range project engages local school communities in the design of their school improvement goals, targets, strategies and measures, providing authentic community engagement.

A specific application of the framework within one community context has focused on enhancing success and achievement for First Nations students. Most data from provincial measures in regard to First Nations students were in the level of concern on the Accountability Pillar. However, jurisdiction leaders acknowledged that such large scale measures

...are not sufficient in painting the whole picture of learner progress. Particularly for First Nation, Métis and Inuit learners, the learning colour palette is incomplete without some consideration of importance of experiential learning via interaction with Elders, other members of community and the lands (Tunison, 2007, p.7).

Provincial data and information gleaned from a literature review were shared with First Nations parents and students, jurisdiction professional staff, and Boards of Education from the two neighbouring First Nations communities. First Nations staff members served as facilitators during the focus groups and community dinners. Key themes emerged, key barriers to success were identified which then led to the articulation of key actions. These formed the basis for development of a plan that included six key outcomes, each with specific measures and targeted strategies.

The plan was and is an ambitious one. One year later, progress is evident not just in the number of new activities and events, but gains are also occurring in student learning outcomes. Two more years remain in the plan. The integration of three year plans with AISI project plans provides a coherent focus that set the stage for success.

This approach requires a new level of commitment and expertise on the part of leaders as they reach out to the community in meaningful ways. Community members gain a deeper sense of ownership than has been evidenced in previous structures where community involvement has often remained at a relatively superficial level. As the number of community members who do not have students in the K – 12 system increases, it is even more critical that school leaders reach out and connect with the community.

The proposed model of pairing newly designed provincial measures with locally determined measures represents a level of accountability that is both robust in terms of alignment with provincial learner outcomes, as well as responsive to the skills and competencies that are valued by the community. It provides for meaningful input and involvement of community members and has the potential to increase the sense of efficacy and ownership among all stakeholders. It is a new way of thinking about public assurance.

One of the major obstacles in adopting such a bold new approach will be the lack of understanding among members of the public at large as to the benefits of this new model of public assurance. The current model is familiar and even though it is becoming more and more evident that the current model will not suffice for a redesigned provincial curriculum framework, change is difficult.

The proposed model of pairing newly designed provincial measures with locally determined measures represents a level of accountability that is both robust in terms of alignment with provincial learner outcomes, as well as responsive to the skills and competencies that are valued by the community.

A clearly articulated implementation plan will be critical to a successful transition, along with samples of the measures in action within Alberta classrooms. The framework for such a model follows. Samples of various measures within a classroom context are available on the AAC website.

PROPOSED IMPLEMENTATION PLAN AND NEXT STEPS

The Alberta Assessment Consortium is well positioned to take on a leadership role in partnership with Alberta Education in the design and implementation of a new provincial assessment program focused on the public assurance model described herein. The Alberta Assessment Consortium

- maintains contact with educational leaders worldwide;
- contributes to assessment literacy within Alberta through development of classroom assessment materials and professional learning materials and opportunities, as well as through action research and inquiry initiatives;
- participates on provincial committees convened by Alberta Education, as well as by a variety of education partners including the Alberta Teachers' Association and Alberta Professional Development; and
- maintains ongoing contact with representatives of post-secondary faculties of education across the province.

The following four key elements are suggested as the framework for an implementation plan.

Phase One: Provincial Measures Assessment Design

As the assessments are designed, clear links to learner outcomes in the English Language Arts and Mathematics Programs of Study must be articulated. Concurrent development must also occur in French. Identifying outcomes at both the grade three and grade four levels will assist teachers in considering the range of achievement levels within the classroom in order to move the learning of all students forward.

Assessments must then be designed and field tested to ensure both the validity and reliability of the assessment measures. Exemplars must be collected and reviewed by teams of teachers in order to establish standards. Exemplars support teachers and students in coming to a shared understanding of various levels of quality within the context of the assessment tasks.

While the assessment rubrics will allow teachers to determine students' current level of performance as measured against the identified learner outcomes, the diagnostic nature of the assessments shifts the focus from providing a grade/level/score to that of identifying students' current strengths and areas of need. This supports teachers as they adjust instruction in order to meet the learning needs of the students.

Phase Two: Professional Learning Support Design

A professional learning framework must be designed to support teachers in preparing to administer the assessments, as well as in using protocols for collaborative teacher scoring of student work. A variety of methods can be employed within the professional learning design, including face to face and technology mediated opportunities. While workshops may form part of the professional learning design, the emphasis will be on job embedded opportunities to enhance teacher assessment practice.

Phase Three: Action Research

Once the assessment measures have been developed and field validated, an action research phase is proposed to test the model within a sampling of school jurisdictions. Potentially four to six jurisdictions would be selected based on the need to represent diverse geographic and student population configurations.

In addition to the student products created during the assessment tasks, additional information from interviews, questionnaires, and focus groups involving teachers, students and parents would inform modifications to the design.

Phase Four: Large Scale Implementation

Large scale implementation would require attention to the seven essential conditions as described in A Guide to Support Implementation: Essential Conditions (Alberta's Education Partners, 2010) a document developed collaboratively by education partners in Alberta.

Successful implementation is possible in environments where education stakeholders are committed to sharing the responsibility for supporting implementation within a culture of learning (p. 1).

A brief description of each essential condition within the context of the proposed model for provincial assessment at the beginning of the grade four year is found in Appendix C.

The implementation of a new model for public assurance has implications for provincial policy and protocols. Some areas of consideration include the following points.

- The current Accountability Pillar would need to be replaced with a new model based on provincial sampling and an emphasis on jurisdiction reporting of student achievement results.
- Resources currently targeted for Provincial Achievement Tests would need to be reallocated to the design of new measures and professional learning opportunities to prepare teachers for administration of the assessments.
- Attention to leadership development in assessment would need to be in place at both the school and jurisdictional levels.
- A communication and public awareness program would need to be developed to ensure that the public and interested stakeholders have information as to the changes, the rationale for those changes, and access to the data.

The implementation of a new model for public assurance has implications for provincial policy and protocols.

Consistent with the province wide public consultation process that has occurred in recent years, the proposed public assurance model has the potential to provide credible evidence of student learning in regard to what Albertans value in their education system. It also has the potential to support teachers in building classroom assessment capacity, providing a focus on learning and growth for all who are involved in the process – students, educators, schools, communities and the province. The shift from external accountability to public assurance will serve to enhance Alberta's position world-wide as a leader in the field of education and learning.

APPENDIX A REVIEW OF LARGE-SCALE TESTING REFORMS IN VARIOUS INTERNATIONAL JURISDICTIONS

As of 2009, England has been, and continues to be, in the middle of a large reform with radical changes to typical high stakes testing for students up to the age of fourteen years. In 2005, in response to concerns that formal testing was not appropriate for very young children, tests at the end of Key Stage 1 (known as SATs, *Standardized Assessment Tests*) were replaced with a system of teacher assessment, informed by tests and tasks and subject to external moderation (Bew, 2011). SATs continue to be given at Key Stage 2 (students aged eleven years) as a key accountability measure, while Key Stage 3 was discontinued due to major problems with test marking in 2008 (Collins, Reiss, & Stobart, 2010). Teacher interviews on Key Stage 2 testing in the area of Science have indicated the welcoming of a movement away from tests, as they believe such a shift would lead to a more complete coverage of the intended curriculum such as allowing for more investigative Science in their teaching (Collins, et al., 2010; Mansell, James, & the ARG, 2009). Continued concerns have also been expressed around evidence of 'teaching to the test' to the extent that the Key Stage 2 SATs may be replaced by tests to be taken on entry to secondary schools. The Key Stage 2 tests are a limited sampling of the National Curriculum; therefore, questions of validity, as well as unreliability due to sampling errors in the choice of test items, have also been raised.

Wales was originally part of the United Kingdom's national testing since its introduction in 1991. Debate began following its introduction as to the validity and impact of such testing and the accountability purposes for which they may be used (Collins, et al. 2010). Key Stage 1 tests (students seven years of age) were abolished in 2001 as assessments were shown to be inadequate for the purposes for which they were intended. Key Stage 2 (students eleven years of age) and Key Stage 3 (students fourteen years of age) tests were withdrawn from national testing in 2005 (Collins, et al. 2010; Mansell, et al. 2009) based on the Daugherty review of assessment policy which found that "their continuation could not be educationally justified" (as cited in Collins, et al. 2010, p. 274). In place of these tests, summative teacher assessments were used, supported by optional test materials at the three levels. The Awdurdod Cymwysterau Cwricwlwm as Asesu Cymru (ACCAC 2004), (the Welsh equivalent to the Qualifications and Curriculum Authority of England) continue to monitor students at Key Stage 2, alongside England, to assess effects of the abolition of these tests (Mansell, et al. 2009), relying on moderated judgments of teachers to summarize attainment at the end of Key Stages 2 and 3.

Northern Ireland has also moved away from 'blanket testing' as it "adds nothing to individual pupils' learning" (Gardner, 2011). Northern Ireland terminated statutory testing in 2005, placing a priority on assessment for learning and teacher assessment (Collins, et al. 2010; Mansell, et al. 2009). Policy changes have moved toward favouring assessment for learning. There has also been a recent change to creation of pupil profiles constructed by teachers.

Scotland reformed its National Qualifications system in 2000 with further changes planned as part of the new Curriculum for Excellence which places a focus on skills, attitudes and capacities for learning, life and work in the 21st century. It no longer collects information on students through national assessments, instead monitoring achievement through sample survey. The reform moved from sampling achievement at the Primary (P3, P5, P7) and Secondary (S2) levels to the Primary (P4, P7) and Secondary (S2) levels, with a focus on literacy and numeracy in alternating years. The sampling approach known as the Scottish Survey of Literacy and Numeracy (SSLN) has been in development since 2009 and consists of a set of written and practical assessments designed to assess the wide range of knowledge, skills, capabilities, and attitudes across the learning outcomes, as well as a questionnaire for students and teachers. The student questionnaire collects information on factors likely to affect learning (attitudes and experience in class) and the teacher questionnaire collects information on how the curriculum has been implemented. Pupil sampling involves students in completing an assessment booklet (online or paper based), participating in an interactive teacher/pupil or group based assessment, and completing the questionnaire. The SSLN provides a national estimate so the number of students required is small (4000 across Scotland). Student selection is random with a goal NOT to assess individual students, rather, to provide information through student achievement on what matters across the curriculum. The survey takes place in three stages and neither schools nor teachers are identified (Mansell, et al. 2009; Scottish Government website, 2012). Policy makers realized that although test results were improving, it was likely that teachers were getting better at rehearsing for the test. With the lessened 'threat' of being identified as individuals, the curriculum isn't narrowed by 'teaching to the test'. The focus and government priority for schools since 2001 has been on assessment for learning.

Finland is known worldwide in its ranking as one of the top educational systems (Booi & Couture, 2011). It does not rely upon external standardized testing to measure the performance of schools nor does it employ a rigorous inspection system of schools and teachers. Instead of test-based accountability, it relies on the expertise and professional accountability of teachers who are knowledgable and committed to their students. Finland publicly recognizes the value of its teachers and trusts their professional judgment in schools, making it a system built on 'trust'. The Finnish government believes a good teacher is someone who is able to help all children progress and grow in a holistic way, and teacher education is designed to ensure educators have well developed curricular knowledge and planning skills, which in turn support school wide improvement efforts. Authorities in education insist that "curriculum, teaching and learning rather than testing should drive teachers' practice in schools" (Sahlberg, 2011). Teachers are considered to be the best judges of how their students are progressing, confirming the responsibility of the school, not external assessors to ensure growth in achievement. Finland (as well as Sweden and Australia) places a great deal of trust in teacher judgments of student achievement, spending their educational dollars on schools, not testing (Booi & Couture, 2011; Wiliam, 2010).

New Zealand is unique in its approach to accountability. Schools have their own governing boards that report to the national Ministry of Education. The Ministry determines resources, establishes curriculum and has authority to close down schools while the Education Review Office (ERO) visits, reviews and reports to the Ministry. National testing is determined by the New Zealand Qualifications Authority and its National Education Monitoring Project (NAEMP) is limited to students in their final years of education (years eleven to thirteen) for purposes of national certification, in decisions for employment or admission to post high school education. Since 1995, key sources for accountability are the Education Review Office and the NAEMP which provide detailed national assessments of knowledge, skills and attitudes of primary and intermediate students at year four (ages eight to nine years) and year eight (ages twelve to thirteen years) through a sampling approach assessing fifteen areas of the national curriculum over four year cycles. The NAEMP utilizes a mix of school based performance assessments and nationally standardized assessments for objectives that cannot be assessed well with standardized tests, such as geometry and statistical investigation. Sixty percent of the items used are performance tasks that are publicly released to provide a regular flow of information on student achievement in the identified key areas. Schools conduct self reviews, while an external ERO review places emphasis on school targets and priorities. According to Professor Terry Crooks at the University of Otago, this approach creates a "sense of negotiation between school and government, rather than an imposition by the Ministry". The focus for schools becomes one that is based on the quality of teaching and learning, enhancing the quality of education (National Center for Fair & Open Testing, 2011).

APPENDIX B RECOMMENDATIONS FOR PROVINCIAL ASSESSMENTS

MEASURES OF LITERACY AND NUMERACY

a. Sources of Reading Assessments

Individualized reading assessments provide teachers with insights into student reading skills and behaviours that are not available through selected response assessments. While listening to the student read, the teacher tracks and records specific reading behaviours, gaining insight into the student's strengths and areas of need. The assessment becomes diagnostic as the teacher links the assessment results to specific instructional strategies to support student learning. When used at both the beginning and the end of the school year, teachers are able to gather evidence of student learning and growth.

Two such reading systems are in wide use in Alberta schools.

Benchmark Assessment System 2, 2nd edition grades 3-8, Irene C. Fountas and Gay Su Pinnell, distributed by Pearson Canada http://www.heinemann.com/products/E02796.aspx

PM Benchmark 1 Reading Assessment, distributed by Nelson Education http://www.gagelearning.com/elementary/langArts/assessment/benchmark.html.

Strengths

- assessment is aligned to what teachers need for quality literacy instruction
- information provided through the assessment identifies strategies being used by the reader, and assists teachers to support struggling readers, who often place an emphasis on guessing when they encounter an unfamiliar word
- provides a wide variety of diagnostic information for teachers to use in their instruction including word miscues, ability to self correct, fluency, and comprehension of and beyond the text
- teachers deepen their understanding of strong literacy practices and are assisted to develop better strategies for teaching reading
- assessments can be recorded through visual or auditory means to use in celebrations of learning, student/parent/ teacher conferences, or for provincial sampling

Considerations

- a one-on-one assessment, therefore the element of time is a factor for teachers in administering the assessment and in recording and interpreting the results
- professional learning support for teachers is a cost factor
- an examination of how the text level gradient aligns to the Alberta Program of Studies would be important in determining common benchmarks if multiple sources are to be used

Reading Logs have been recommended as an assessment for determining basic understanding, inference, and analysis. Students read a predetermined number of books in identified areas and provide their thoughts, based on a series of questions developed either locally or provincially. As reading is a process that includes the demonstration of comprehension and an interpretation of text, reading logs assist students to make connections between parts of a text, between texts and experiences, and to extend and apply their understanding. A recommendation would be for students to read a minimum of twenty-five texts or book equivalents, including fiction and non-fiction, across a variety of contemporary and traditional literature. Students would then provide evidence of their comprehension through a group of common questions. The written sample would then be assessed from the perspective of reading comprehension, not students' writing ability. It could also provide information on student attitudes about reading.

Strengths

- students choose books of interest as they practice the skill of learning to read
- growth in the learner is focused as students write and receive feedback from teachers through the use of questioning and specific comments
- rubrics connected to curricular outcomes in areas such as exploring thoughts, ideas, feelings and experiences; and expressing ideas and preferences may be used to assess this form of response

Considerations

- some students may only read within narrow areas of literature or respond only superficially
- variance in teacher grading may occur without the availability of exemplars
- · observable reading behaviours (e.g., fluency, decoding) could not be assessed through this method

b. Possible Source for Writing Assessment

The **Highest Level of Achievement Test (HLAT) Teacher Resource** is a quality writing assessment developed by Edmonton Public School Board and used across the district. The HLAT is intended to provide a measure for determining student achievement relative to the graded curriculum in English Language Arts. Student writing is assessed as to grade level and performance level. The tests have been designed to assess a wide genre of writing over time.

Development of the HLAT involves determining a prompt. The prompt is then field tested by districts in Alberta who provide samples of student writing. A team of assessment and language arts consultants, working within the district, assess the samples against outcome criteria. Exemplars are chosen that can be clearly linked to the criteria, and a brief explanation for why a given sample meets a specific performance level is written. Information is provided to teachers through a resource manual that includes instructions on administration, literacy resources that support the prompt and a variety of student samples at grade level that meet different levels of performance. The resource is designed to be used by teachers in marking their own students' writing. Prompts from previous years are available to be used with students in a formative capacity. The task also includes a self reflection piece by the student.

The HLAT provides an opportunity for focused teacher professional learning in assessing student writing. Teams of teachers can be brought together to discuss what makes for an excellent, proficient, adequate or limited piece of writing to ensure reliability in marking. As teachers dialogue as to grade level and level of performance, they determine strategies that can be used to move students from level to level. Development of a similar provincial HLAT resource could include professional development for teachers working together to determine exemplars for use across the province.

Strengths

- · exemplars are developed through collaborative processes that provide professional learning in assessment practices
- schools mark together collaboratively as a way of developing consistency and accuracy when judging writing at the specific grade level, an excellent form of professional learning in developing quality writers
- the prompt is common across all grade levels, a factor in reducing development costs
- specific information about student strengths and areas for growth can be used to determine instructional strategies

Considerations

- writing is a first draft product completed within a specific time frame
- time in marking collaboratively may be a factor for teachers

c. Performance Assessment Possibilities

Performance assessments are open-ended tasks that assess learner understanding through extended constructed responses and performances. They provide a more contextual measure of higher level thinking skills and hence the ability to assess multi-literacies (e.g., technology use, collaborative skills). They can be developed provincially or through teacher collaboration, providing a professional development component.

Examples of performance tasks in many subject areas are currently available on the AAC website for use by teachers. These tasks provide an opportunity for teachers to assess skills that are more difficult to measure through traditional means, by allowing students to apply their knowledge, skills and understanding in new and innovative ways. Possible examples include the following.

- create a digital Photostory demonstrating understanding of the 'rules of recess', meeting outcomes in language arts, health and ICT
- use Moviemaker to create a trailer that uses effective visual and oral communication and demonstrates respect for various cultures, meeting outcomes in language arts and ICT
- solve a problem related to an authentic purchase for the class to demonstrate graphing, organization and communication skills in numeracy

Strengths

- can be employed to address outcomes that are more difficult to measure such as speaking, listening, viewing, and representing
- can be collaboratively scored, increasing assessment capacity of teachers and leaders within a school or jurisdiction through development of common understanding of standards
- require rubrics and exemplars, so protocols for reliable scoring can be developed
- artifacts of performance tasks can be recorded and managed through the use of digital portfolios for sharing with parents or used as exemplars for teachers

Considerations

- teacher comfort with the use of technologies as a possible tool must be considered, as well as equitable access to technologies in order to support a variety of programs for students to use
- more time consuming to administer and to score, although some of this time is more focused on learning than with traditional paper/pencil assessments

d. Sources for Self Reflections

Student Questionnaires provide a way to assess a variety of student attitudes and engagement in areas such as enjoyment in reading and writing; and confidence in reading, writing, speaking, listening, viewing and representing. Questionnaires help to build a broad picture of the student, collecting information on factors that are likely to affect learning, such as attitudes toward and experiences in educational settings.

Student questionnaires such as the *Learning Bar* survey [http://www.thelearningbar.com] have been used by Alberta schools. A sampling similar to one used by the Scottish Government [http://www.scotland.gov.uk/Publications/2010/02/22135127/0] could also be used as a model and developed to address the Alberta context.

Strengths

- involves students in monitoring their own learning, assisting them to become less dependent on the teacher
- provides information that can be used by the teacher to plan instruction and meet specific student needs

- can be designed to provide information about a variety of behaviours and attitudes including reading frequency both in and out of school, preferred learning style, external support for schooling, interests and so forth
- information can be collected and accessed digitally to attain a class, school or district picture
- honors the rights of students to express themselves and be heard

Considerations

- accuracy of data collected is dependent on how students feel in reporting truthfully about their learning and behaviours
- time to administer survey and available access to technologies

Learning Logs can be used at the end of a lesson, or at specific times in the year; students complete a learning log entry (either written or through a voice thread) using a variety of prompts. Information provided by the students provides the teacher with insights as to how students learn and where they need more support. Information can also be provided on the growth a student has shown from one assignment to the next. Samples of prompts could include:

- The easiest part for me was...
- The hardest part was...
- Today I learned...
- One thing that interested me was...
- One thing I could use more help with is...
- What engaged me the most in this lesson was...
- In the past I had difficulty with...now I have learned...

Strengths

- involves students in monitoring their own learning, assisting them to become less dependent on the teacher
- · provides information that can be used by the teacher to plan instruction and meet specific student needs
- students are provided with opportunities to regulate cognitive resources, emotions and actions to assist in attaining their learning goals
- students learn the importance of participating in their learning, which increases motivation and guides them through key metacognitive processes

Considerations

• self reflection is a skill that must be learned and given time and attention in the classroom

Learning Portfolios have been in use by teachers for years. However, the majority have been performance portfolios, a collection of the 'best' evidence of learning. In contrast, learning portfolios provide a look at work over time and allow the student, as well as other stakeholders, to see how they have improved and to reflect and set goals for future learning. An advantage of learning portfolios is the focus on improvement, so all students see themselves as having the ability to grow. Learning portfolios involve the students in seeing growth in their abilities as incremental rather than as fixed. When used for learning, the teacher organizes a set time for students to access the portfolio, adding or reflecting on pieces within the portfolio to demonstrate evidence that meets a specific outcome. They can be kept digitally to archive growth in a variety of areas.

Strengths

- involves students in monitoring their own learning, assisting them to become less dependent on the teacher
- provides information that can be used by the teacher to plan instruction and meet specific student needs
- students are provided with opportunities to regulate cognitive resources, emotions and actions to assist in attaining their learning goals
- students learn the importance of participating in their learning which has the potential to increase motivation; help them monitor what they know and can do; and help them to understand more about how they learn
- supports the development of autonomous learners

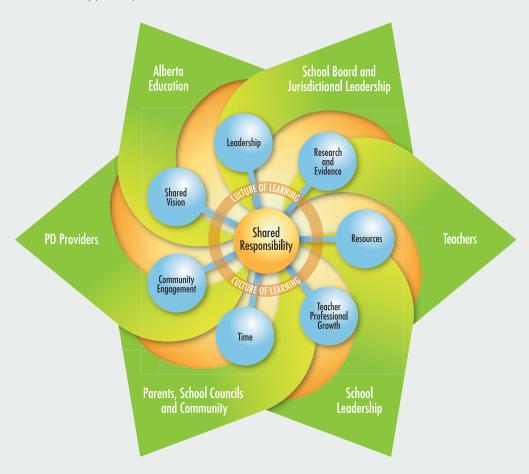
Considerations

- involves a time commitment from the teacher on an on-going basis
- may involve an issue of space if not kept digitally

APPENDIX C ESSENTIAL CONDITIONS TO SUPPORT IMPLEMENTATION

The following seven conditions have been identified as being essential to supporting effective implementation. Brief descriptions have been provided that link each essential condition with the proposed model for public assurance, based on student assessments at the beginning of the grade four year.

Figure 2
Essential Conditions to Support Implementation



Shared Vision

Stakeholders share an understanding of and commitment to the intended outcomes.

All education stakeholders have a vested interest in a successful public education system. A redesigned provincial assessment model consistent with a competency focused curricular framework has the potential to provide the desired level of public assurance as well as to enhance classroom assessment practices.

Leadership

Leaders at all levels have the capacity to champion the shift from the current reality to the intended outcomes.

The proposed model requires the commitment of jurisdiction and school based leadership to the process of building assessment capacity, with an understanding that the process of building capacity takes place over time. Leaders must be prepared to support teachers in this process, both conceptually and through appropriate allocation of resources.

Research and Evidence

Current research, evidence, and lessons learned inform implementation decisions.

A worldwide research base confirms the positive impact of formative assessment practices on student learning. The proposed model is both based on research and has the potential to provide action research evidence to inform future design of provincial assessments.

Resources

Human resources, materials, funding, and infrastructure are in place to realize the intended outcomes.

Estimates of the current costs of administering the provincial testing program are in the millions of dollars annually. While the initial costs of restructuring the provincial assessment model would likely not realize any cost savings in the short term, the long term re-allocation of these funds towards enhancing teacher assessment capacity has the potential to realize benefits thus far unattained.

Teacher Professional Growth

Teacher knowledge, skills and attributes are enhanced through ongoing professional learning.

Job embedded teacher professional learning is foundational to the design of the model. Consistent with the research base on effective professional learning designs, such a model would support teachers in the administration of the assessments, using collaborative scoring protocols including the use of student exemplars, and in using the results of the assessments to adjust instruction in order to meet the learning needs of students. By design, it seeks to build classroom assessment capacity over time.

Time

Time is provided to support implementation.

The significant change in focus from an accountability framework to a public assurance framework will require time to ensure effective implementation. Appropriate time must be allocated at each stage, beginning with the development and field validation of the measures and development of the professional development support; testing of the prototype and subsequent adjustments informed by the prototype; and communication with educators, parents, students and the public.

Community Engagement

Parents, school councils, students, community members, businesses, industry and post-secondary institutions are partners in supporting implementation.

The proposed model addresses the needs expressed by Albertans to ensure the education system prepares students to acquire the competencies for engaged thinkers and ethical citizens with an entrepreneurial spirit (Alberta Education, 2011). A multifaceted assessment approach, including performance based assessment, provides opportunities for students to be deeply engaged in their learning and to experience assessment tasks within real world contexts. An advisory committee with representation from various stakeholder groups would be an important element of the implementation plan.

REFERENCES

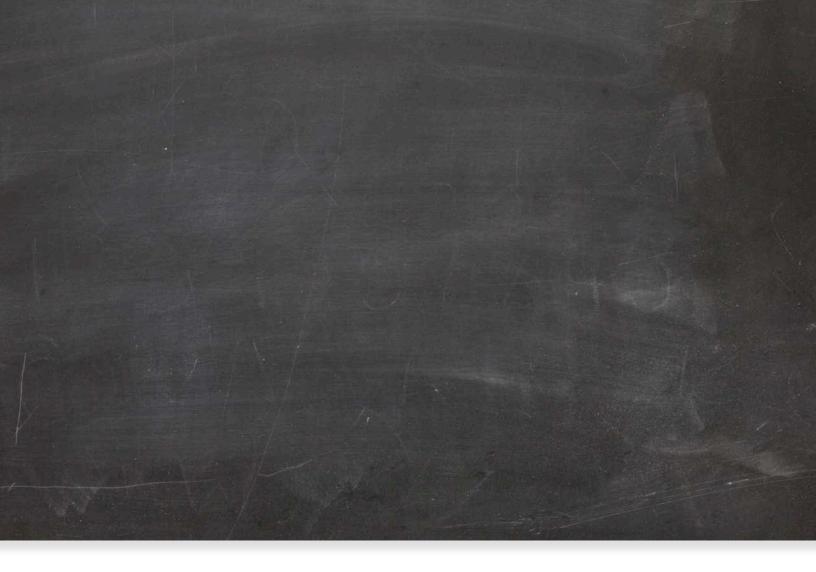
- Adrian, L. (2012.) AISI Scoop (January ed.). Retrieved February 29, 2012, from, http://education.alberta.ca/media/6648257/scoop_january_2012.pdf.
- Assessment Reform Group (ARG). (2002). Testing, motivation and learning. Pamphlet, University of Cambridge Faculty of Education. Retrieved December 28, 2011 from, http://www.assessment-reform-group.org.uk.
- Alberta Education. (2010). Inspiring education: A dialogue with Albertans. Edmonton, AB: Alberta Education.
- Alberta Education. (2011). Framework for student learning. Retrieved March 1, 2012, from http://education.alberta.ca/department/ipr/curriculum.aspx.
- Alberta Education. (2012). Ten point plan for education. Retrieved February 18, 2012, from http://alberta.ca/NewsFrame.cfm?ReleaseID=/acn/201201/31784C87E013F-07B2-09FF-E745E76F1B9DFC1F.html.
- Alberta Education. (2012a). AISI cycle 5 handbook 2012–2015. Retrieved March 26, 2012, from http://education.alberta.ca/media/6657808/handbook for aisi cycle 5 projects.pdf.
- Alberta Education. (2012b). School community engagement rubric: Moving from vision to action. Retrieved March 26, 2012, from http://education.alberta.ca/media/6674139/school community engagement rubric 2012.pdf.
- Alberta's Education Partners. (2010). A guide to support implementation: Essential conditions. Edmonton, AB: Alberta's Education Partners.
- Awdurdod Cymwysterau Cwricwlwm as Asesu Cymru. (2004). Retrieved December 28, 2011, from http://www.smspmaesteg.org.uk/Equal%20Opportunities%20and%20Diversity%20in%20the%20Curriculum.pdf.
- Bew, Lord P. (2011). Independent review of key stage 2 testing, assessment and accountability, final report, as written for the Department of Education. Retrieved December 28, 2011, from https://www.education.gov.uk/publications/standard/publicationDetail/Page1/DFE-00068-2011.
- Booi, L., & Couture, J.C. (2011). Testing, testing. What Alberta can learn from Finland about standardization and the role of the teacher. *Alberta Views*, 7, (pp. 28-32).
- Collins, S., Reiss, M., & Stobart, G. (2010). What happens when high-stakes testing stops? Teachers' perceptions of the impact of compulsory national testing on science of 11-year-olds in England and its abolition in Wales. *Assessment in Education: Principle, Policy & Practice*, 17(3), (pp. 273-286).
- Costa, A. L. & Kallick, B. (2003). Assessment strategies for self-directed learning. Thousand Oaks: Corwin Press.
- Davies, A. (2011). Making classroom assessment work, 3rd edition. Courtney, B.C.: Connections Publishing.
- Doecke, B., Kostogriz, A., & Illesca, B. (2010). Seeing "things" differently: Recognition, ethics, praxis. *English teaching practice and critique* 9(2), (pp. 81-98).
- École Providence School. (2010). École providence school annual education plan, 2010. Retrieved September 15, 2011, from http://www.hfcrd.ab.ca/Providence/2010-2011%20EPS%20Ed%20Plan%20&%20AERR.pdf.
- Fountas, I.C., & Pinnell, G.S. (2007). Benchmark assessment system 2, 2nd ed., grades 3-8. Toronto: Pearson Canada.
- Fullan, M. (2011). The moral imperative realized. Workshop session presented for the College of Alberta School Superintendents in January of 2012. Calgary, AB.
- Gardner, J., (2011). Personal correspondence by email September 2011. Vice-President of the British Educational Research Association; Visiting Professor, University of Oxford Centre for Educational Assessment; School of Education, Queen's University, Belfast, Northern Ireland.

- Hamilton, L. (2010-2011). Testing what has been taught; Helpful, high-quality assessments tart with a strong curriculum. *American educator* (Winter), (pp. 47-52).
- Hargreaves, A. (2009). In praise of accountability. In K. Gariepy, B.L. Spencer, & J-C. Couture (Eds.), Educational accountability: Professional voices from the field. Rotterdam: Sense Publishers.
- Jones, B., & Egley, R. (2004). Voices from the frontlines: Teachers' perceptions of high stakes testing. In *Education policy analysis archives*(EPAA) 12(39). Retrieved December 28, 2011, from http://www.eric.ed.gov/PDFS/EJ853506.pdf.
- Livingstone Range School Division. (2011). Enhancing first nation success & achievement annual results report 2010-2011 & Three Year Plan 2010-2013. Retrieved March 26, 2012, from http://www.lrsd.ab.ca/ppp/plans%20and%20Reports/Documents/First%20Nation%20Success%20Plan%20and%20AERR%20Nov%202011.pdf.
- Mansell, W., James, M., & the Assessment Reform Group. (2009). Assessment in schools. Fit for purpose?

 A commentary by the Teaching and learning research programme. London: Economics and Social Research Council, Teaching and Learning Research Programme.
- McElligott, D. (2012, May 10). Premier Redford, *Alberta at Noon*. Edmonton, Canada: Canadian Broadcasting Corporation.
- McTighe, J. & Wiggins, G. (2010). Measuring what matters: Part 1 The case for an assessment overhaul. Hope Newsletter (January, 2011).
- Murgatroyd, S. & Sahlberg, P. (2010). Accountability, learning and the teacher Looking at real learning first. Edmonton: Alberta Teachers' Association.
- Murgatroyd, S. (2011). Rethinking education: Learning and the new renaissance. Edmonton, Canada: Future Think Press.
- National Center for Fair and Open Testing. (2011). New Zealand accountability model. Retrieved December 28, 2012, from http://fairtest.org.
- Sahlberg, P. (2011). Lessons from Finland. In The professional educator American educator (Summer), (pp. 34-39).
- Sands, A. (2012, May 15). New education minister outlines his top priorities. Edmonton Journal, A4.
- Scottish Government, Riaghaltas na h-Alba (2012). Retrieved December 28, 2011, from http://www.scotland.gov.uk/Topics/Statistics/Browse/School-Education/SSLN.
- Sutton, R. (2005). *Embedding assessment for learning*. Workshop session presented for Edmonton Public Schools in the Fall of 2005. Edmonton, AB.
- Tunison, S. (2007). Aboriginal learning: A review of current metrics of success. University of Saskatchewan, Aboriginal Education Research Centre, Saskatoon, Saskatchewan and First Nations and Adult Higher Education Consortium, Calgary, Alberta. Retrieved March 29, 2012, from www.fnahec.org.
- Volante, L. (2004). Teaching to the test: What every educator and policy-maker should know. Canadian journal of educational administration & policy 35. Retrieved December 28, 2011, from http://www.umanitoba.ca/publications/cjeap/articles/volante.html.
- Wiliam, D. (2010). Assessing Achievement at the end of key stage 2. Prepared for the Department of Education, U.K. in response to Key Stage 2 testing. Personal correspondence by email December 10, 2011.
- Wiliam, D. (2011). Embedded formative assessment. Bloomington, IN: Solution Tree Press.

BIBLIOGRAPHY

- Alberta Assessment Consortium. (2006). Refocus, 2nd ed. Edmonton, AB: Alberta Assessment Consortium.
- Alberta Assessment Consortium. (2008). *Building classroom assessment capacity to enhance student success.*Project Charter written for Alberta Education to fulfill Conditional Grant A.
- Alberta Assessment Consortium. (2011). Moving forward: A vision of assessment in Alberta in the 21st century A draft discussion document. Available online: http://www.aac.ab.ca.
- Black, P., & Wiliam, D. (2006). The reliability of assessments. In J. Gardner (Ed.), Assessment and learning (pp. 9-26). London: Sage.
- Harlen, W. (2004). A systematic review of the evidence of the reliability and validity of assessment by teachers for summative purposes. In J. Gardner (Ed.), *What's working in high schools*. Courtney, B.C.: Connections Publishing.
- Jaafa, S.B. & Earl, L. (2008). Comparing performance based accountability models: A Canadian example. Canadian journal of education, 31 (3), (pp. 697-726). Retrieved December 28, 2011, from http://www.eric.ed.gov/PDFS/EJ809266.pdf.
- Pink, D. (2006). A whole new mind: Why right-brainers will rule the future. New York: Riverhead Trade.
- Webber, C., Aitken, N., Lupart, J., & Scott, S. (2009). *The Alberta student assessment study: Final report.*Alberta Education. Retrieved December 28, 2011, from http://education.alberta.ca/department/ipr.aspx.





#700, 11010 - 142 Street Edmonton, Alberta Canada T5N 2R1

www.aac.ab.ca

