Chapter 4
Student Assessment Policy and Practice in Alberta: An Assessment for Learning

Jim Brandon and Marsi Quarin-Wright

Introduction

This chapter examines student assessment policy and practice within the context of the 2009 *Alberta Student Assessment Study* (ASAS) (Webber et al. 2009). At a time of unprecedented interest in student assessment, we explore the degree to which educational policy and practice in the Canadian province of Alberta reflect the best available research evidence in each of the ASAS's four focus areas: (1) quality teaching, (2) educational leadership, (3) professional learning, and (4) education policy. From our perspectives as practice-based researchers, we aim to gage aspects of contemporary assessment policy and practice through an *assessment for learning* approach for the purpose of enhancing educational policy and practice.

The chapter is divided into four parts. A general overview of our foundational assumptions is followed by an explanation of the nine research-informed standards featured in our conceptual framework. These standards are then used in two *formative assessments.* We first look at assessment practice before applying the standards to selected Alberta provincial policies. Discussion and implications are framed by the four major goals of the ASAS:

1. Define optimal assessment theory and practice relative to:
   (a) Curricular learning outcomes and performance standards and reporting of levels of achievement within grade.

J. Brandon (✉)
University of Calgary
e-mail: jbrandon@ucalgwy.ca

M. Quarin-Wright
Foothills School Division, High River, AB, Canada
e-mail: quarin-wrightm@fsd38.ab.ca

C.F. Webber and J.L. Lupart (eds.), *Leading Student Assessment*,
Studies in Educational Leadership 15, DOI 10.1007/978-94-007-1727-5_4,
© Springer Science+Business Media B.V. 2012

(b) How external tests and classroom-based assessment of student achievement can be optimally used to inform decisions regarding student program needs and, when aggregated upwards, decisions at the school, jurisdiction, and provincial levels of decision-making relating to improving learning opportunities for students.

2. Describe how educational leadership can be strengthened to facilitate more effective classroom assessment and accurate and meaningful reporting of assessment information, including grade level of achievement, to parents and to Alberta Education.
3. Identify professional development models that are needed to build and enhance capacity in the area of classroom assessment (of, for, and as learning) in Alberta schools.
4. Consider the question, "where do we go from here" in developing a holistic framework for classroom assessment for the province and provide recommendations based on what is learned from the preceding questions and areas of inquiry. Our conclusion conveys four student assessment paradoxes in response to the ASAS Goal Four question: "where do we go from here in developing a holistic framework for classroom assessment for the province?" These paradoxes suggest ways in which educational policymakers and practitioners can begin to move past solidified perspectives and contested views to improve policy and practice.

Conceptual Framework

This section highlights the conceptual underpinnings and assumptions that guide this assessment for learning. Figure 4.1 is an idealized representation of relationships among educational research, policy, and practice. The outer circle indicates that (a) education research should inform both policy and practice, (b) research-informed policy should influence practice, and (c) the contextualized realities of educational practice should inform both research and policy. The pyramid within the

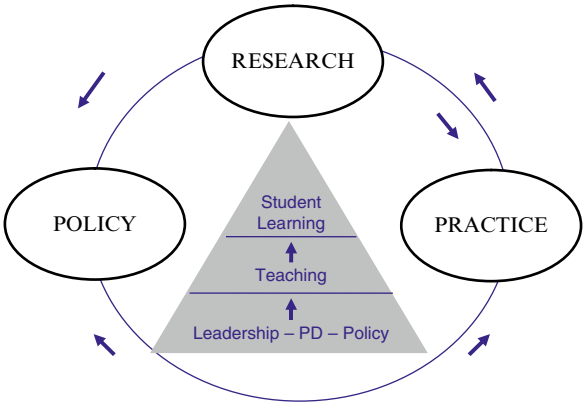

Fig. 4.1 Conceptual framework

Fig. 4.2 Assessments for policy and practice learning

circle suggests ways in which policy, leadership, and professional learning indirectly impact student learning. Our assessment for learning approach is portrayed in Fig. 4.2. Areas in which research-informed standards have been developed are shown in the upper rectangle. The degree to which current practice and selected provincial policies reflect the best available evidence in selected areas of policy and practice is illustrated in the oval at the bottom.

Our conceptual framework models effective assessment practice. We first provide images of research-informed policy and practice. These images are then used as standards for our assessment for learning. The design parallels effective classroom assessment wherein clearly delineated learning targets or learner outcomes are the starting points for student assessment.

Assessment Standards

Effective assessment starts with a clear vision of what is to be assessed. As Stiggins and Chappuis (2005) make evident, "We can't dependably assess that which we have not defined" (p. 2). Each of the assessment standards presented in this section is designed to create a clear picture of evidence-based policy or practice that will subsequently be used in the assessment for learning sections below. We begin with the general theoretical stance that shapes the development and use of our standards. In separate subsections, we then move into descriptions of the standards and the evidence upon which each is based.

The advice of two scholars has been helpful in the formulation of our standards. Marzano advocated for "lean and mean standards that are specific and non-redundant" (as cited in Scherer 2001, p. 4). Darling-Hammond (1994) observed:

> Standards are useful only to the extent that teachers can use them to build their own knowledge and understanding of what helps students learn… it is the process of using a set of images about teaching and learning to deepen one's own understanding and that of the teachers and students with whom we work that makes standards useful in any way (p. 10).

Table 4.1 A research-informed image of quality teaching

Quality teaching as situated, collective expertise-in-action that
• Is contingent on the dynamic interplay of content, teacher, learner, and context
• Involves professional commitment to collegial practice and reflection over time
• Accesses scientific and artistic pedagogic tools in a fluent, seamless, holistic, and constantly evolving pattern of practice

Just as Darling-Hammond sees merit in "using a set of images… to deepen understanding" in teaching, it is our desire to use research-informed images to deepen our understanding of student assessment policy and practice.

Standard One: Quality Teaching as Situated, Collective Expertise-in-Action

In our view, any discussion of classroom assessment needs to understand assessment as an aspect of teaching. Our first assessment standard is expressed through the image of *quality teaching as situated, collective expertise-in-action*. This notion is based on our reading of leading authors in this field (i.e., Bennett and Rolheiser 2001; Danielson 1996; Darling-Hammond 1994, 1999, 2001; Darling-Hammond and Sclan 1992; Fullan 2003; Marzano et al. 2001; Organization for Economic Co-operation and Development 1994; Stigler and Hiebert 1999; Stronge 2002).

That quality teaching is situated in a given context is noted in several studies. Darling-Hammond and Sclan (1992) promoted conceptions of teaching as "informed judgment" rather than as "mastery of simple routines." Emphasis on the former takes "a reflective teaching orientation stimulated by attention to teachers' individual contexts and felt needs" rather than emphasizing the "production of specific teacher behaviors thought to represent effective teaching" (Darling-Hammond and Sclan 1992, p. 15) (Table 4.1).

The idea of *expertise-in-action* is a significant component of quality teaching. Fullan (2003) indicates that the future of teaching rests in the notion of "through informed professional judgment" (p. 7). Stronge (2002) supported this view of dynamic expertise and noted that the "teacher must have sufficient knowledge of content, of pedagogy, of context, and of students to appreciate the intricacies that are bound up in the teaching and learning process" (p. 63). Lieberman and Miller (1999) painted a picture of pedagogy that goes beyond a mere set of technical skills by constructing good teaching as "a complex array of values, knowledge, experience, intuition, and commitment to improve" (p. 63). Darling-Hammond (1994) noted that "teaching is intense activity, that it requires juggling of subject matter, cognitive goals, social goals; management of time, materials and equipment; and the needs and responses of individual students" (p. 18). These observations also underline the situated aspect of the standard.

Table 4.2 A research-informed image of formative assessment

Formative assessment as a key aspect of teaching and learning that provides students with clear pictures of progress and how to improve – during the learning process by
• Utilizing ongoing descriptive and encouraging feedback to support learning
• Fostering student involvement in, reflection on, and ownership of the learning process
• Providing feedback to inform instruction

The collective aspect of the standard's conception of quality teaching has only recently appeared in the literature. Bennett and Rolheiser (2001) call for extended professional learning leading to "collectively conscious instructional intelligence" or expert teacher "intuition informed by experience combined with the experience and research of others" (p. 46). Fullan (2003) advocated "collective deliberations focusing on continuous improvement" (p. 6). Similar research support for the fundamental importance of the collective improvement of teaching is found in the work of Elmore (2000, 2002, 2004), Marzano et al. (2001), and Stigler and Hiebert (1999).

The standard development approach employed is similar to the processes utilized by educational organizations, such as the Committee on Standards for Educational Evaluation (Sanders 1997). Our Standard One was first written in draft form based upon a review of applicable research (Brandon 2005). The standard was then provided to others for review, and alterations were made based upon the feedback received. Evolutionary refinements were made through reflection and consideration of new insights gleaned from ongoing reading in this content area. The same approach was used in the development of Standards Five, Eight, and Nine below.

Standard Two: Formative Assessment as Generative and Informative Teaching

Our second assessment standard is expressed through the image of *formative assessment as generative and informative teaching.* The image's first indicator underlines the importance of providing students with a clear picture of progress and how to improve during the learning process. Students need to know where they are in relation to where they are going in their learning. This is achieved through the teacher's skillful use of ongoing, descriptive, and encouraging feedback. In our conception, formative assessment is a key aspect of generative and informative teaching that fosters student involvement in, reflection on, and ownership of the learning process.

A strong body of research supports the concepts expressed through this standard. Black and Wiliam (1998), Chappuis and Chappuis (2007), Chappuis et al. (2005), Costa and Kallick (2004), Marzano (2006), and Tomlinson (2007) are among the chief works consulted in the construction of this image of formative assessment (Table 4.2).

Table 4.3 A research-informed image of summative assessment, grading, and reporting

Summative assessment, grading, and reporting that provide students and their parents with clear pictures of achievement in relation to learning outcomes in the Program of Studies or in an Individual Program Plan – at the end of a learning episode
• Provides consistent, accurate, and outcome-referenced descriptions of learning
• Based upon informed professional judgment using varied assessment tools to show best available evidence of learning
• Fosters student involvement in, reflection on, and ownership of the learning process

We first drafted the standard in early January of 2008 based upon our review of the assessment literature noted above. Early drafts were discussed and refined through conversation with several school-based leadership colleagues and with those in the Foothills School Division System Leadership Team. The standard later evolved into one of three components of a College of Alberta School Superintendents (CASS) position on Research-Informed Student Assessment. Minor adjustments were made in consultation with the CASS Provincial Executive in early April 2008. The three assessment standards were circulated to all 258 CASS members and were adopted by at the Annual General Meeting on April 25, 2008.

Standard Three: Summative Assessment, Grading, and Reporting as Consistent, Accurate, and Outcome – Referenced Descriptions of Learning

This third assessment standard is expressed through the image of *summative assessment, grading, and reporting as consistent, accurate, and outcome – referenced descriptions of learning.* Similar to the evidence on formative assessment, it is imperative to provide students and their parents with clear pictures of achievement in relation to learning outcomes in the Program of Studies or in an Individual Program Plan. In the case of summative assessment, the picture of achievement in relation to learning outcomes is communicated (by letter, symbol, number, or statement) at the end of a unit, term, or school year. Consistent and accurate professional judgments must be informed through the application of appropriate and varied assessment tools to show the best available evidence of learning. The literature supports focusing on what the student knows and is able to do in relation to clearly identified learning targets. Such an approach fosters student involvement in, reflection on, and ownership of the learning process (Table 4.3).

The educational literature and research vein upon which this standard is based is represented by the work of several authors (i.e., Davies 2000; Marzano 2006; O'Connor 2002, 2007; Reeves 2007). The standard drafting process for Standard Three was the same as for Standards Two and Four.

Table 4.4 A research-informed image of external assessment

External assessments as complementary outcome – referenced descriptions of learning that provide snapshots of achievement in relation to a significant portion of the outcomes of the Program of Studies

- Provide data over time to aid policymaking and curriculum development
- Provide data over time to be used in combination with classroom, school, and jurisdiction data to inform longer-term instructional, school, and system improvement planning
- One of many assessment tools to inform professional judgment in relation to summative assessment, grading, and reporting

Table 4.5 Percent of Program of Studies outcomes covered in provincial examinations

Course	Grade 3	Grade 6	Grade 9	Grade 12
English language arts	70	66	69	75
Social studies		74	83	
Social studies 30				75–80
Social studies 33				80–85
Mathematics	95	83	88	85–95
Science/sciences		70	73	65

Source: Alberta Education, Accountability and Reporting Division (November 2007)

Standard Four: External Assessment as Complementary Outcome – Referenced Descriptions of Learning

Our final assessment standard is expressed through the image of *external assessment as complimentary outcome – referenced descriptions of learning* (Table 4.4). External assessments, such as the Alberta Diploma Examinations and the Provincial Achievement Tests, provide snapshots of achievement in relation to a significant portion of the outcomes in the Program of Studies (see Table 4.5). The image's first indicator underlines that a primary benefit of external exam is to policymakers and curriculum developers. Results over time provide data that are useful for policy development and program review. A second value of external assessment is the provision of data to inform longer-term instructional, school, and system improvement planning. The final descriptor calls for use of external assessment as one additional tool at the disposal of the teacher to inform professional judgment in relation to summative assessment, grading, and reporting.

A number of authors were studied in the development of this standard (e.g., Davies 2000; Marzano 2006; O'Connor 2002, 2007; Reeves 2007). As in the drafting of Standards Two and Three, this evidence-based image was adopted by CASS at its Annual General Meeting on April 25, 2008, following considerable discussion at the local and provincial levels.

Table 4.6 A research-informed image of professional learning

Professional learning as coherent, incremental capacity building that
• Expects and nurtures staff learning within a community of professional practice
• Integrates ongoing opportunities for reflection, professional dialogue, and continuous pedagogic learning in – or directly related to – the school setting
• Gradually improves student learning, increases teacher efficacy, and builds school capacity

Standard Five: Professional Learning as Coherent, Incremental Capacity Building

To examine the "professional development models that are needed to build and enhance capacity in the area of classroom assessment" as outlined in the fourth goal of the ASAS, we utilize the image of *professional learning as coherent, incremental capacity building.* Several works were instrumental in developing this standard (i.e., Elmore 2002; Fullan 2001a, b, 2003; Guskey 2000, 2003; Lieberman and Miller 1999; Richardson 2003; Sparks 2002; Stigler and Hiebert 1999) (Table 4.6).

One major idea in this image of research-informed professional learning is coherence. The standard's first indictor is that the learning of all staff and students is both expected and nurtured in the context of community of professional practice. Learning communities help in the development of coherence, as observed by Lieberman and Miller (1999), in that they are "organizational forms that provide for support and pressure" (p. 72). "Professional learning is most powerful," they further elaborated, "when it occurs as a result of one's being a member of a group of colleagues who struggle together to plan for a given group of students, replacing the traditional isolation of teachers from one to another" (p. 62). Fullan (2001a) called for program coherence to avoid "too many disconnected, episodic, piecemeal, superficially adorned projects" (p. 109).

Incremental staff learning is the second main component. Sparks (2002) indicated that "powerful forms of professional development engage teachers in the continuous improvement of their teaching" through "training, coaching, critical friends and other reflective processes" (pp. 10–14). Fullan (2001b) supported "learning in the setting where you work, or learning in context" because it has the "greatest payoff" (p. 126). Lieberman and Miller's (1999) *growth-in-practice* idea reflects the need to rebalance the content of teacher learning to include more of what they refer to as "'inside knowledge' – by teaching and picking up ideas from fellow teachers and trying them out in their classroom," what Schon (1987) called "an epistemology of practice that takes fuller account of the competence practitioners… display in situations of uncertainty, complexity, uniqueness, and conflict" (p. 63). Stigler and Hiebert (1999) came to a similar conclusion that improvement will be continual, gradual, and incremental because "teaching is a system deeply embedded in the surrounding culture of schools" (p. 132). Incremental staff learning is integrative and provides ongoing opportunities for teacher reflection, professional dialogue, and continuous pedagogic learning in – or directly related to – the school setting.

Table 4.7 A research-informed school leadership standard – the PQPG

The school principal is an accomplished teacher who provides quality leadership in the provision of optimum learning and development for all students in the school
1. Fostering effective relationships
2. Embodying visionary leadership
3. Leading a learning community
4. Providing instructional leadership
5. Developing and facilitating leadership
6. Managing school operations and resources
7. Understanding and responding to the larger societal context

The third key concept within this standard is capacity building. Fullan (2003) observed, "it takes capacity to build capacity" (p. 7). To Glickman et al. (2001), we must "return wisdom, power, and control to both individuals and the collective staff in order for them to become true professionals" (p. 56). Guskey's (2000) insight was that "change in teacher attitudes and beliefs occur only after teachers changed their practices and they begin to see the results of these changes in terms of student outcomes" (p. 68). Stigler and Hiebert (1999) made a similar point:

> Teaching is a system built from all elements of the local context: teacher, students, curriculum, and so on. Improving the system requires taking all of these elements into account… Teaching is unlikely to improve through researchers' developing innovations in one place and then prescribing them for everyone. Innovations can be spread around the country, but only by trying them out and adjusting them again and again as they encounter different kinds of classrooms (p. 133).

The final indicator ties staff learning to three key results: the gradual improvement of student learning, an increase in teacher efficacy, and enhanced school capacity. This standard was originated through the same approach described in the development of Standards One, Eight, and Nine (Brandon 2008).

Standard Six: Principal Quality Practice Guideline

The next standard used in our analysis is Alberta's *Principal Quality Practice Guideline* (PQPG). Table 4.7 presents the standard's seven leadership dimensions. The descriptors, which provide detailed expectations related to each leadership dimensions, are detailed in the complete document (Alberta Education 2008).

The PQPG development process was led by an active stakeholder advisory committee, which referred to the research-informed positions on the principalship developed by the Alberta School Boards Association, Alberta Teachers' Association, CASS, Interstate School Leaders Licensure Consortium, and Alberta Home and School Councils' Association. The committee, which began its work in the spring of 2005, found that the research literature and Alberta Education partners agreed that individuals designated as principals require a broad repertoire of competencies

Table 4.8 A research-informed system leadership standard – the CPS

The CASS member is an accomplished leader and teacher who ensures that each student is provided with the opportunity to achieve optimum learning
1. Visionary leadership
2. Instructional leadership
3. Human resource leadership
4. Ethical leadership
5. Effective relationships
6. Organizational leadership and management
7. External influences on education
8. Chief executive and chief education officer leadership

to successfully fulfill their complex and critical roles within the education system. A first draft of the practice standard was completed in early 2006.

In the fall of 2006, focus group meetings were held across the province to provide opportunities for sharing information and receiving feedback regarding the *draft Principal Quality Practice Standard*. Approximately 170 participants representing the various stakeholders were involved in the focus group meetings. Written feedback from the small groups and individual participants was submitted and has subsequently been compiled for the consideration of the Stakeholder Advisory Committee.

Standard Seven: College of Alberta School Superintendents Practice Standard

The *CASS Practice Standard* (CPS) is the second leadership standard used in our analysis. Table 4.8 presents the *standard* and its eight leadership dimensions. The descriptors, which provide detailed expectations related to each leadership dimension, are detailed in the complete document (College of Alberta School Superintendents 2008).

The CPS was designed through a process that mirrored the PQPG development process in a number of ways. A consultant and a stakeholder advisory committee developed a first draft early in the fall of 2007, following a review of research and consideration of standards developed in similar jurisdictions. The draft was then taken to the general membership in November of 2007, with changes coming back to the committee for consideration in early 2008.

The consultant then took this second draft to each of the five CASS zones for further discussion and refinement. Approximately 200 of the 250 CASS members contributed to its development. After a final meeting of the committee in the spring of 2008, the standard underwent a legal review that has led to small wording changes. The CPS was adopted at a Special General Meeting in November 2008.

Table 4.9 A research-informed standard – quality education policy content

Education policy content informed by
• Research
• Exemplary practice
• Policy learning

Standard Eight: Quality Education Policy Content

The basic premise of research-informed policy content is that quality education policy occurs when policymakers – through ongoing analysis of the social, political, and educational context – design and enact policy content based on the prevailing research consensus in each of the policy's major conceptual areas (Brandon 2005).

Is it reasonable to expect all important education policies to meet similar research-based tests in content areas where a scholarly consensus exists? It would be naïve to think that research would ever be the sole determinant of a government's policy agenda. However, it is reasonable, in our judgment, to expect that significant policy choices should, at the very least, be justified on the basis of recognized education research. In cases where policy choices are being made in the advance of an existing research consensus, policy impacts should be carefully studied from the early stages of implementation. To the greatest extent possible, the content of significant education policy should be founded on current research. In policy content areas where no clear research consensus exists, policy impacts should be carefully studied from the early stages to enable research-based refinements along the way (Table 4.9).

Policies based on research have increased chances of succeeding, as do those that are based on evidence of positive results in similar jurisdictions through policy learning or lesson-drawing (Bennett and Howlett 1992; Hall 1993; Howlett and Ramesh 1995; Rose 1993). As Rose (1993) noted, politicians learn from experience, whereas policy analysts more frequently take their lessons from formalized research sources (p. 19). An additional element supported by our research review is that effective policy designs should ensure that policies are well matched to their political context and that key values in the policy align with those prevalent in wider society.

Standard Nine: Quality Education Policy Process

Quality education policy occurs when policymakers – through ongoing analysis of the social, political, and educational context – design and enact research-based proposals through processes characterized by dialogic adoption and implementation as learning (Brandon 2005). Though policy development is a nuanced, iterative, and complex political undertaking, four stages often characterize quality education policy development: (a) informed design, (b) dialogic adoption, (c) implementation

Table 4.10 A research-informed standard – quality education policy process

Education policy development through

- Informed design *that*
  - Pursues modest goals with a real chance to improve student learning
  - Is based on research or evidence of positive results in similar settings
  - Is part of comprehensive and coherent plan
  - Matches proposals to political and social contexts
- Dialogic adoption *that*
  - Provides for public input and open debate through authentic consultation
  - Yields modified and improved proposals through consultation, collaboration, or partnering
- Implementation as learning *that*
  - Provides ongoing support and capacity building in the work setting
  - Is viewed as ongoing policy learning with opportunities for refinement based on experiences in the field
- Meaningful outcomes *that*
  - Are clearly and coherently explained in the design phase
  - Are framed so that the benefits to students are clearly evident

as learning, and (d) meaningful outcomes. These four phases coincide closely with the four stages in Levin's (2001) policy process model. Among others, the following eight studies have significantly influenced the thinking that yielded this policy process standard: Earl et al. (2002), Elmore (2000, 2002, 2004), Fullan (2001a), Hightower (2002), Leithwood et al. (2002), Levin (2001), Levin and Wiens (2003), and Wilson et al. (2001). Each of the phases is now briefly explained.

The informed design research consulted suggests that policy reform is more likely to achieve its intentions when goals are clearly focused on those things that are likely to yield desirable student outcomes. By focusing energy and resources on achievable targets, policymakers demonstrate thoughtful stewardship of available resources. Attention to design clarity enhances opportunities for practitioners to develop understanding and ownership as they translate the policy into action (Leithwood et al. 2002, p. 9).

Dialogic adoption calls for adoption strategies that provide opportunities for public input and open debate through authentic consultation processes. Quality policy processes should actually yield modified and improved proposals through stakeholder consultation and collaboration. Several of the studies consulted noted that governments are becoming more aware of the benefits of public debate and the incorporation of a variety of voices in the policy process. An open, inclusive approach to policy adoption can build legitimacy for the proposal. As Hightower (2002) observed, "building support for the change is crucial, in a politically volatile situation, as is building professional support" (p. 23) (Table 4.10).

The standard's *implementation as learning* notion indicates that an effective implementation plan should provide ongoing support and capacity building for those who are expected to translate policy intentions into practice. A learning orientation to implementation taps into educator motivation and heightens

efficacy. Another indicator of effective implementation is that the process of implementation itself is used as a basis for ongoing policy learning. Refinements to practice are made in the field based on the actual implementation experiences of educators.

A substantive research base supports these ideas about implementation. Fullan (2001b) described a "recent remarkable convergence of theories, knowledge bases, ideas, and strategies that help us confront complex problems… a new mind-set – a framework for thinking about and leading complex change more powerfully than ever before" (p. 3). Capacity building implementation strategies are founded on the understanding that policymakers must pay more attention to the perspectives of those in the field. In the end, it is the "street-level" commitment and actions of school-based educators that determine the success or failure of policy initiatives (Lipsky 1980).

The standard's conception of *meaningful outcomes* asserts that in the final analysis, educational policy changes should be judged on the student learning outcomes they generate. It is, however, very difficult to precisely determine effects of any individual policy, whether considering student outcomes, education system outcomes, or social outcomes. What seems to be significant in the outcome phase is the notion that greater clarity and coherence can be developed when meaningful outcomes are spelled out in the design phase. Also worth noting in this short treatment of the construct of meaningful outcomes is this recurring finding: Educators will be more willing to translate policy into practice when they believe that the policy will make a positive difference for students (Earl et al. 2002; Elmore 2000, 2002, 2004; Fullan 2001a; Hightower 2002; Leithwood et al. 2002; Levin 2001; Levin and Wiens 2003).

Assessments for Learning – Improving Teaching, Leading, and Professional Learning

The research-informed assessment standards presented in section "Assessment standards" were designed to create clear pictures of quality student assessment policy or practice. These standards will be used in sections "Assessments for learning – Improving teaching, leading, and professional learning" and "Assessments for learning – Improving provincial education policy" to reflect upon the state of student assessment practice and to analyze selected policies, standards, and professional development structures in the Canadian province of Alberta – our assessments for policy and practice learning.

In each case, elements of the formative assessment method we utilized are explained first. Next, a summary table conveying our assessment is presented. Explanatory remarks related to the assessment are then provided to conclude each subsection. The first four assessments are designed to provide information to enhance efforts to develop research-informed teaching and leadership practice in the case of Representative School Division (RSD). Assessment Five analyzes three

Alberta professional learning structures within the context of RSD. The same subsection assesses two additional professional development institutions from an overall provincial perspective.

Our assessments are not meant to be definitively defensible assessments for policy and practice learning. Rather, they represent attempts to nest our reflections on current policy and practice within a conceptual framework that links to what we know about quality assessment. Our analyses of practice in RSD are qualitative judgments based upon observations and conversations over a 6-month period in 2008. As these data were translated onto assessment tables to communicate descriptions of performance, the initial assessments were reviewed and verified by two separate groups of educational leaders. A group of six school administrators provided initial feedback to adjust our assessments. Additional refinements were then made through conversation with the nine members of the RSD's System Leadership Team. Our views were further influenced by the writing of three members of the System Leadership Team, who recently reported on several aspects of RSD administrator and teacher growth in evidence-based assessment practice.

Representative School Division is a medium-size suburban school system. It serves approximately 7,000 students, with a teaching staff of nearly 400 in 25 schools (including 7 that may be classified as alternative schools). The system does quite well on measurable provincial outcomes, with recently documented strengths in high school achievement, staff learning, provision of a safe and caring environment, and offering students a strong and broad program of studies. A three-pronged approach to capacity building seems to be serving the district well: (a) persistent attention to shared instructional leadership, (b) a learning community emphasis through the Alberta Initiative of School Improvement (AISI), and (c) ongoing implementation of a district-wide learning coaching program. Learning more about research-informed student assessment has been a focus for the district through these means for the past 2 years.

Assessment One: Quality Teaching as Situated, Collective Expertise-in-Action

This subsection's two assessments for learning reflect on current teaching and leadership practice in RSD in relation to our research-informed standard of *quality teaching as situated, collective expertise-in-action.* Our descriptive feedback utilizes the same three-point scale used in teacher and administrator evaluations in the province. Practice is described as exceeding, meeting, or not meeting the descriptor's expectations. Reflections are based to a large extent on our general observations and self-assessing conversations with school and system administrators. Additional insights were provided through reflecting on 24 administrative evaluations and over 100 teacher evaluations written or reviewed over the past 4 years.

We believe that Table 4.11 presents a clear picture of pedagogic and leadership progress in relation to the standard of quality teaching as situated, collective

Table 4.11 Assessment One

Evidence of quality teaching in representative school division	Teaching practice	Leadership support
Quality teaching as situated, collective expertise-in-action that		
• Is contingent on the dynamic interplay of content, teacher, learner, and context	+	+
• Involves professional commitment to collegial practice and reflection over time	=	+
• Accesses scientific and artistic pedagogic tools in a fluent, seamless, holistic, and constantly evolving pattern of practice.	=	=

Scale: + exceeding; = meeting; – not meeting

expertise-in-action. RSD teachers and administrators exceed the expectations expressed through the first indicator: Quality teaching is contingent on the dynamic interplay of content, teacher, learner, and context. Observed teachers demonstrated strengths in adjusting outcome-focused teaching in response to the unique needs of their learners within the distinct contexts of their schools. Evaluations written by administrators and conversations related to classroom visits with school administrators indicate that school leaders understand the descriptor and possess the necessary supervisory skill set.

Teachers and school administrators meet the expectations expressed in the other two descriptors. In the case of *professional commitment to collegial practice and reflection over time*, our analysis suggests that leaders have a strong understanding and appreciation of the impact of working through a learning community approach. Many representative leaders have focused on this area of growth through their professional learning plans over the past few years. Teachers, particularly those who have engaged in AISI shared leadership work or have served as learning coaches, are demonstrating increasing strength and commitment to this approach to collective practice. Both groups of professionals intuitively understand and demonstrate a fluent, seamless, holistic, and constantly evolving pattern of practice. We see evidence of movement toward more fully articulated conceptualization of both scientific and artistic pedagogic tools in both cases.

Assessment Two: Formative Assessment as Generative and Informative Teaching

Our second assessment considers teaching and administrative practice in relation to the image of *formative assessment as generative and informative teaching.* The descriptive feedback summarized in Table 4.12 utilizes a four-point scale common to many assessment rubrics. Practice is described along a continuum: exemplary, skilled,

Table 4.12 Assessment Two

Evidence of research-informed formative assessment in representative school division	Teaching practice	Leadership support
Formative assessment as a key aspect of teaching and learning that provides students with clear pictures of progress and how to improve – during the learning process by		
• Utilizing ongoing descriptive and encouraging feedback to support learning	D	S
• Fostering student involvement in, reflection on, and ownership of the learning process	D	D
• Providing feedback to inform instruction	D	S

Scale: *E* exemplary; *S* skilled; *D* developing; *P* partial

developing, and partial. Judgments in this case are based on our general observations along with conversations with teachers, school, and system administrators.

Evidence indicates that the provision of research-based professional development opportunities to school administrators has increased the understanding and use of formative assessment in RSD schools. In particular, administrators have shown a skilled level in using feedback to encourage learning and providing feedback to inform instructional practice and are providing professional development opportunities to staff in this area. Further support is evident through the work of well-trained learning coaches and the use of external educational consultants. Several principals have taken staff members to national and international conferences to deepen and expand teacher understanding.

The Division should continue to develop administrator and teacher understanding and skill in the area of fostering student involvement in, reflection on, and ownership of the learning process. Continued professional learning about the value of student ownership of learning is advised.

Assessment Three: Summative Assessment, Grading, and Reporting as Consistent and Accurate Outcome – Referenced Descriptions of Learning

The image of *summative assessment, grading, and reporting as consistent, accurate outcome – referenced descriptions of learning* guides this third assessment. The assessments of RSD teaching and administrative practice summarized in Table 4.13 use the same four-point scale used in Assessment Three. Comments on practices in comparison to the standard are again based on general observations along with conversations with teachers, school, and system administrators.

Representative School Division has engaged administrators and teachers in professional learning about how to best inform students and their parents with a clear picture of achievement in relation to the learning outcomes or individual program

Table 4.13 Assessment Three

Evidence of research-informed summative assessment, grading, and reporting in representative school division	Teaching practice	Leadership support
Summative assessment, grading, and reporting that provide students and their parents with clear pictures of achievement in relation to learning outcomes in the Program of Studies or in an Individual Program Plan – at the end of a learning episode		
• Provides consistent, accurate, and outcome-referenced descriptions of learning	D	S
• Based upon informed professional judgment using varied assessment tools to show best available evidence of learning	D	S
• Fosters student involvement in, reflection on, and ownership of the learning process	D	S

Scale: *E* exemplary; *S* skilled; *D* developing; *P* partial

plan at the end of a learning episode. This has helped to improve summative assessment practices. Evidence indicates tighter alignment with the Program of Studies and increasingly clear communication with parents and students about what is being taught and evaluated.

Assessment Four: External Assessment as Complementary Outcome – Referenced Descriptions of Learning

Two different groups are the objects of this subsection's assessments for learning. We reflect on current provincial and system leadership practice in relation to our research-informed standard of *external assessment as complementary outcome – referenced descriptions of learning.* Our descriptive feedback utilizes the same four-point scale used in Assessments Two and Three. Judgments are based primarily on our observations and conversations with system educational leaders and Alberta Education management personnel in a variety of settings over an extended period.

A high degree of skilled knowledge and practice is evident in senior educational leadership ranks across the province. There is widespread appreciation that external assessments, such as the Alberta Diploma Examinations and the Provincial Achievement Tests, provide snapshots of achievement in relation to a significant portion of the outcomes in the Program of Studies (see Table 4.5). Both system education leaders and ministry managers work from strong operational understanding of the image's first indicator. There is agreement that the primary benefit of external exams is to policymakers and curriculum developers. Results over time provide data that are useful for policy development and program review. Another agreed upon value of external assessment is in the provision of data that inform longer-term instructional, school, and system improvement planning when used in

Table 4.14 Assessment Four

Evidence of research-informed external assessment	System leaders	Alberta education
External assessments as complementary outcome – referenced descriptions of learning that provide snapshots of achievement in relation to a significant portion of the outcomes of the Program of Studies		
• Provide data over time to aid policymaking and curriculum development	S	S
• Provide data over time to be used in combination with classroom, school, and jurisdiction data to inform longer-term instructional, school, and system improvement planning	S	S
• One of many assessment tools to inform professional judgment in relation to summative assessment, grading, and reporting	S	S

Scale: *E* exemplary; *S* skilled; *D* developing; *P* partial

combination with classroom, school, and jurisdiction data. While both groups understand this, the government position does seem to emphasize the external exam result in itself as the "gold standard" when determining how well a school or jurisdiction is doing in a particular area (Table 4.14).

On the final descriptor, both groups recognize and are beginning to reap the benefits of using external assessment as one additional tool at the disposal of the teacher to inform professional judgment in relation to summative assessment, grading, and reporting. If we are indeed "on the threshold of an exciting new educational context" (Brandon 2008, p. 9), then the climate may be ripe for more open dialogue about what the provincial examinations can and cannot offer in the way of student assessment insights. Productive results are more likely if all parties steer clear of rigidly entrenched ideological or organizational positions.

Assessment Five: Professional Learning as Coherent, Incremental Capacity Building

Our fifth assessment focuses on the third goal of ASAS: "to identify professional development models to build capacity in assessment (of, for, and as) learning." The evidence indicates that this aim can be achieved by optimizing current structures. Individual professional growth planning, AISI learning community work, and school-based professional development in RSD stand up well in terms of the expectations of the research-informed standard of *teacher growth as coherent, incremental capacity building*. The descriptive feedback summarized in Table 4.15 utilizes another four-point scale common to many assessment rubrics. Practice is described along a frequency continuum: consistently, usually, occasionally, and seldom.

Table 4.15 Assessment Five

Evidence of research-informed professional learning	IPGP	SBPD	AISI	AAC	RPDC
Teacher growth as coherent, incremental capacity building that					
• Expects and nurtures staff learning within a community of professional practice	4	4	4	3	3
• Integrates ongoing opportunities for reflection, professional dialogue, and continuous pedagogic learning in – or directly related to – the school setting	3	4	4	2	2
• Gradually improves student learning, increases teacher efficacy, and builds school capacity	3	3	3	3	3

Note: Scale = consistently (*4*); usually (*3*); occasionally (*2*); and seldom (*1*)
IPGP Individual Professional Growth Plan; *SBPD* school-based professional development; *AISI* Alberta Initiative for School Improvement; *AAC* Alberta Assessment Consortium; *RPDC* Regional Professional Development Consortium

Professional growth blossoms when there is coherent support for educator learning at all levels: individual, school, district, and region/province. In RSD, professional learning opportunities are provided in frequent increments. At least 1 day a month is dedicated to professional development for staff at the school or district level. Staff members also have the opportunity to learn through their Individual Professional Growth Plans, where each staff member is expected to choose a topic, learn from it, and apply it to his or her professional practice. These are expected to align with school or divisional objectives. Success is based on a combination of the willingness of the staff member to engage in this process and administrator skill in engagement through reflective dialogue. School-based professional development generally follows the school's self-identified AISI plan. School-based professional development sessions vary from administrator or teacher leader presentations of research-informed practice to the use of guest speakers or more focused staff work in professional learning communities.

District-wide initiatives on assessment are provided to administrators and teachers on an ongoing basis. Administrators and learning coaches facilitate many additional staff learning opportunities in the area of student assessment. Educators in the district also connect with regional or provincial professional development activities to scaffold individual, school, or division professional learning. Specifically, the district provides support for a team of teachers to work with the Alberta Assessment Consortium to become assessment specialists. A relationship with the Regional Professional Development Consortium is fostered so that teachers may attend Regional Professional Development Consortium–offered courses that align with personal, school, or divisional professional learning.

Each approach to professional development has the potential to engage educators in reflection and ongoing improvement of assessment practice. Research suggests that locally provided professional development has a greater impact on teaching practice.

At the same time, both consortia offer quality experiences from a distance. Hence, the consortia receive a lower assessment on Table 4.15. Distant opportunities do not as easily lend themselves to integrating deep reflection, professional dialogue, and continuous pedagogic learning.

Assessments for Learning – Improving Provincial Education Policy

While Assessments Six and Seven focus on the third ASAS goal (strengthening leadership), our analysis also sheds light on Goal Four (policy recommendations). In this subsection and the next three, our reflections are based upon our observations, document analysis, related literature, and conversations with key provincial leaders. The scales used in these assessments mirror the scales used in Alberta Education's color-coded *Accountability Pillar Report.*

Assessment Six: Policy Content – Leadership Quality Standards

A quick glance at Table 4.16 makes it clear that both the draft *Principal Quality Practice Standard* and the draft *CASS Practice Standard* fare well as research-based policy positions. When each is reviewed for its content in relation to our four teaching-related standards, they do well. The excellence that we awarded to the PQPG and the CPS in relation to our evidence-based image of quality teaching starts from the fact that in each case, the standard statement calls for the leader to be "an accomplished teacher" who focuses leadership efforts on providing all students with optimum learning opportunities. In each case, leaders are guided by leadership dimensions, such as instructional leadership, effective relationships, visionary leadership, leading a learning community, and external influences on education. Within these dimensions are numerous descriptors that underline the importance of understanding and demonstrating quality teaching in the two leadership roles.

Neither leadership standard rates excellence in any of the other four assessment content areas. They both rate *good* for the reflection of research-informed content with respect to our images of formative, summative, and external assessment. Each draft standard acknowledges the importance of assessment. The PQPG requires principals to "ensure that student assessment and evaluation practices throughout the school are fair, appropriate and balanced" (Alberta Education 2008, p. 6). The CPS goes a little further in expecting that:

- CASS members ensure alignment of teaching and student assessment with the provincial curriculum.
- Student learning is assessed, evaluated, and reported using a fair, appropriate, and balanced program of multiple indicators and sources of evidence.
- Student assessment is used to shape and inform instruction (CASS 2008, p. 2).

Table 4.16 Assessment Six

Evidence of quality policy content	Principal quality practice standard	CASS practice standard
Content informed by research on		
• Quality teaching	E	E
• Formative assessment	G	G
• Summative assessment	G	G
• External assessment	G	G

Scale: *E* excellent; *G* good; *A* acceptable; *I* issue; *C* concern

That the PQPG and CPS provide strong conceptual starting points for addressing the ASAS goal of strengthening educational leadership at the system and school levels to enhance classroom assessment practice is already well recognized across Alberta. Provincial and regional educational organizations (e.g., CASS, the Alberta Teachers' Association, the Centre for Leadership in Learning), Alberta universities, and several school jurisdictions are already using the seven PQPG leadership dimensions as the learning outcome framework for their leadership development or graduate programs in educational administration. At the district level, there is a growing number of systems that are employing variations of the PQPG to anchor their administrative growth, supervision, and evaluation programs. CASS plans to use the CPS to build a system of modules as requirements for professional certification.

Assessment Seven: Policy Process – Leadership Quality Standards

As models of quality policy development, the draft PQPG and draft CPS can also provide insights in support of the fourth ASAS goal that focuses on the development of a holistic policy framework for classroom assessment for the province. Both leadership standards were developed through processes described in Assessment Standards Six and Seven (section "Assessment standards"). Table 4.17 communicates our very positive assessment of these policy development processes used to generate each standard. They are good examples of what can happen when policymakers design and enact evidence-based proposals through processes characterized by dialogic adoption and implementation as learning. Our color-coded assessments reveal excellent processes across the board in both cases.

Assessment Eight: Policy Content – Selected Alberta Education Policies

Attention now turns to the ASAS mandate to develop policy recommendations to shape classroom assessment in the province. Observations, document analysis, related studies, and conversations with key provincial leaders are used to examine

Table 4.17 Assessment Seven

Evidence of quality policy process	Principal quality practice standard	CASS practice standard
Policy development through informed design		
• Modest goals focused on student learning	E	E
• Based on research or results in similar settings	E	E
• Is part of comprehensive and coherent plan	E	E
• Matches political and social contexts	E	E
Dialogic adoption		
• Public input, open debate, and authentic consultation	E	E
• Modified and improved proposals through consultation processes	E	E
Implementation as learning		
• Provides ongoing support and capacity building		
• Viewed as ongoing policy learning with field-based refinements		
Meaningful outcomes		
• Clearly and coherently explained in design phase	E	E
• Framed so that benefits to students are clearly evident	E	E

Scale: *E* excellent; *G* good; *A* acceptable; *I* issue; *C* concern

Table 4.18 Assessment Eight

Evidence of quality policy content	TQS 2006	GLA	CAA	PAT	PDE	Preliminary 6/9 PAT Reporting
Research-informed content						
• Formative assessment	G		A			
• Summative assessment	G	G	A	G	G	C
• External assessment	E	A	G	E	E	C

Scale: *E* excellent; *G* good; *A* acceptable; *I* issue; *C* concern; *CAA* Computer-Adapted Assessment Program; *PDE* Provincial Diploma Examination

six existing provincial policy positions. Our assessments in relation to the research-informed standard for quality policy content are portrayed in Table 4.18 through the government's Accountability Pillar Reporting color scheme. Comments on each of the policies follow.

Our assessments of the content of five of the six policies are quite favorable. The 2006 supplement to the *Teaching Quality Standard* (Alberta Education 1997) titled *Effective Student Assessment and Evaluation in the Classroom* (TQS 2006) is

research-based and compares favorably to all three of our assessment standards. Similarly, the idea of *Grade Level of Achievement* (GLA) initiative is well founded on the summative evaluation research. Claims to support the reporting of GLA to the province are not as well supported. Evidence to support the content of two yearly provincial examinations as good summative assessments and excellent examples of external assessments is well founded.

Though we see the promise of the Computer Adaptive Assessment program, its content is strongest in the realm of external and summative assessment at this point in its development. While the potential to support teacher judgments in the area of formative assessment is recognized, we rate it as *acceptable* at this time. We find that Alberta Education's requirement for Preliminary Reporting of multiple choice portions of the grades 6 and 9 Provincial Achievement Test (PRPAT) results directly to parents for the first time at the end of the current year is not supported by the research. In fact, the PRPAT initiative has the potential to further undermine teacher support for the PAT program, in our view.

Assessment Nine: Policy Process – Selected Alberta Education Policies

Analysis of the six selected policies in relation to our quality policy process standard is presented in Table 4.19. A mixed picture of Alberta provincial policymaking emerges. From a positive viewpoint, there is strong evidence of excellent policymaking in each of the four development components: informed design, dialogic adoption, implementation as learning, and meaningful outcomes. On the other hand, the evidence reveals issues and concerns that should spur reflection and policy learning on the part of Alberta Education. Discussion of policy processes is now presented in relation to the four policy development phases.

Three of the selected policies provide examples of excellent policy development through informed design. The 2006 supplement to the *Teaching Quality Standard* (Alberta Education 1997) titled *Effective Student Assessment and Evaluation in the Classroom* (TQS 2006) and two provincial examination programs have modest goals that focus on student learning, are research-based, and are part of a coherent plan that fits the provincial education context to a large measure. These are excellent examples of informed design.

While the GLA is founded on summative assessment research, there is little evidence to support the requirement of schools to report of GLA to the province. The fact that this initiative has been such a political hot-button issue is attributable to the perception that it is part of a larger unwelcome program that could provide one more point of exposure for schools and districts through expanded Accountability Pillar Reporting. Though the GLA remains a highly contested policy, it may be that it was "a little too far ahead of the curve" for the context into which it was introduced. It may yet endure as a heroic first step toward heralding the virtues of teacher judgment as the new "gold standard" in student assessment.

Table 4.19 Assessment Nine

Evidence of quality policy process	TQS	GLA	CAA	PAT	PDE	Preliminary 6/9 PAT reporting
Policy development through:						
Informed design						
• Modest goals focused on student learning	E	G	I	E	E	C
• Based on research or results in similar settings	E	A	A	E	E	C
• Is part of comprehensive and coherent plan	E	A	I	E	E	C
• Matches political and social contexts	E	I	I	E	E	C
Dialogic adoption						
• Public input, open debate, and authentic consultation	A	I	I	E	E	C
• Modified and improved proposals through consultation processes	A	A	I	E	E	C
Implementation as learning						
• Provides ongoing support and capacity building	I	E	A	G	G	C
• Viewed as ongoing policy learning with refinement based on field experiences	I	E	A	G	G	C
Meaningful outcomes						
• Clearly and coherently explained in design phase	E	A	A	A	A	C
• Framed so that benefits to students are clearly evident	E	A	A	A	A	C

Scale: *E* excellent; *G* good; *A* acceptable; *I* issue; *C* concern

Quite good examples of dialogic adoption and implementation as learning can be found in processes historically used to develop and implement provincial examinations. Considerable stakeholder consultation has characterized the development of these examinations over the years. Refinements have been made to their content and administration based on field experiences. TQS 2006 provides a good policy example of developmental processes that clearly and coherently explain the meaningful outcome in the design phase and that are framed so that the benefits to student learning are clearly evident.

The jury is still out on the design of the Computer Adaptive Assessment program. It may well end up having very positive impacts on student learning, but the large amount of government money assigned to the project at a time of resource scarcity and the outsourcing of the project to private corporate interests have undermined stakeholder support for the goals of the project.

The Preliminary Reporting of multiple choice portions of the grades 6 and 9 PATs is a good example of how not to develop provincial policy. This hastily implemented

and little discussed initiative has the potential to undermine public confidence in teacher assessment. It is an example of a policy that runs strongly against both the policy content and policy process research. This initiative needs to be withdrawn as soon as possible.

Improving Student Assessment in Alberta – Paradoxes for Progress

The assessments for learning presented in this chapter are intended to generate further reflection, dialogue, and inquiry to spur continuing progress in student assessment policy and practice in Alberta. We now focus on the larger question posed in ASAS Goal Four, "where do we go from here in developing a holistic framework for classroom assessment for the province?" Four paradoxes arise from our analysis. As Deal and Peterson (1994) explained, each paradox represents "a seemingly contradictory situation or statement that runs counter to common sense and yet appears to be true" (p. 41). The four paradoxes are "to be embraced and creatively addressed, not to be seen as an either-or choice" (Deal and Peterson 1994, p. 9). In our view, the way forward is not a simple linear path. Rather, the pathway to further progress is through embracing the nuanced complexity of evidence-informed policy and practice.

Paradox One: While formative assessment has tremendous promise for improving student learning and is enthusiastically embraced by classroom teachers, summative assessment, grading, and reporting must be given equal attention in improving student assessment practice.

Research-informed assessment is not an either-or proposition. We need to ensure that students and their parents receive accurate and consistent descriptions of progress in relation to provincial or individual program plan outcomes. The challenge is how to do this in such a way that student confidence and ownership of learning developed through quality formative assessment are not undermined.

Paradox Two: There is considerable research evidence to inform classroom assessment practice. It is important to use such evidence to help educators to develop informed professional judgment rather than to impose informed prescriptions to govern practice.

There is an underlying misconception whereby professional development should be based on having practical ideas that can be used in the classroom as soon as possible. This notion must be challenged. It is important for teachers to understand the research behind good practice and to strive for deeper understanding. It is only when teachers fully embrace the "why" of good practice that truly professional teaching will become widespread.

Paradox Three: There are legitimate concerns about the misuses of external assessment; nevertheless, external assessments are a necessary mechanism for building confidence in the provincial school system.

Alberta students perform well on provincial, national, and international assessments. And public confidence in the Alberta education system is quite high despite legitimate educator concerns about media-reported competitive ranking of schools by a neoconservative political organization. Alberta Education should continue to participate in and report on large-scale assessments while following the lead of the Ontario government in negotiating media agreements to lessen the public profile given to the misuses of external assessment results.

Paradox Four: Alberta Education and key provincial stakeholders hold sharply divergent views on approaches to student assessment. In order to sustain momentum in improving student assessment in the province, movement toward a greater consensus is necessary.

The ASAS is providing an exemplary public forum through which a new provincial consensus direction for student assessment is possible. Forward progress is contingent upon Alberta Education, the Alberta Teachers Association, and other stakeholder groups striving for common ground through the dispassionate consideration of the best available evidence. All parties must engage in solution-focused dialogue to capitalize on the opportunities presented at this time.

References

Alberta Education. (1997). *Ministerial order (#016/97) teaching quality standard applicable to the provision of basic education in Alberta*. Edmonton, AB: Alberta Government.

Alberta Education. (2007). *Presentation to the Program Standards and Accountability Stakeholder Advisory Committee*. November 2007, Edmonton, AB.

Alberta Education. (2008). *Draft principal quality practice: Successful school leadership in Alberta*. Edmonton, AB: Alberta Government.

Bennett, B., & Rolheiser, C. (2001). *Beyond Monet: The artful science of instructional integration*. Toronto, ON: Bookation.

Bennett, C., & Howlett, M. (1992). The lessons of learning: Reconciling theories of policy learning and policy change. *Policy Sciences, 25*(3), 275–294.

Black, P., & Wiliam, D. (1998). Inside the black box: Raising standards through classroom assessment. *Phi Delta Kappan, 80*, 139–148.

Brandon, J. (2005). *A standards-based assessment of Alberta's teacher growth, supervision and evaluation policy*. Unpublished doctoral dissertation, University of Calgary, Calgary.

Brandon, J. (2008). Assessment for leading in the new Alberta context. *The CASS connection* (Spring, 2008), 9.

Chappuis, C., & Chappuis, J. (2007). The best value in formative assessment. *Educational leadership, 65*(4), 14–18.

Chappuis, C., Stiggins, R., Arter, J., & Chappuis, J. (2005). *Assessment for learning: An action guide for school leaders*. Portland, OR: Assessment Training Institute.

College of Alberta School Superintendents. (2008). *Draft CASS practice standard*. Edmonton, AB: College of Alberta School Superintendents.

Costa, A., & Kallick, B. (2004). Launching self-directed learners. *Educational leadership, 62*(3), 51–55.

Danielson, C. (1996). *Enhancing professional practice: A framework for teaching*. Alexandria, VA: Association for supervision and Curriculum Development.

Darling-Hammond, L. (1994). *Standards for teachers*. New York: American Association of Colleges for Teacher Education.

Darling-Hammond, L. (1999). *Reshaping teaching policy, preparation and practice: Influences of the National Board for Teaching Standards*. New York: American Association of Colleges for Teacher Education.

Darling-Hammond, L. (2001). Standard setting in teaching: Changes in licensing, certification, and assessment. In V. Richardson (Ed.), *Handbook of research on teaching* (pp. 751–776). Washington, DC: American Educational Research Association.

Darling-Hammond, L., & Sclan, E. (1992). Policy and supervision. In C. Glickman (Ed.), *Supervision in transition* (pp. 7–29). Alexandria, VA: Association for Supervision and Curriculum Development.

Davies, A. (2000). *Making classroom assessment work*. Courtenay, BC: Connections Publishing.

Deal, T., & Peterson, K. (1994). *The leadership paradox: Balancing logic and artistry in schools*. New York: Jossey-Bass Publishers.

Earl, L., Watson, N., & Torrance, N. (2002). *Front row seats: What we've learned from the national literacy and numeracy strategies in England.* Paper presented at the Annual Meeting of the American Educational Research Association in New Orleans.

Elmore, R. (2000). *Building a new structure for school leadership. A policy paper prepared*. Washington, DC: The Albert Shanker Institute.

Elmore, R. (2002). *Bridging the gap between standards and achievement: The imperative for professional development in education. A policy paper prepared*. Washington, DC: The Albert Shanker Institute.

Elmore, R. (2004). *School reform from the inside out: Policy, practice and performance*. Cambridge, MA: Harvard Education Press.

Fullan, M. (2001a). *The new meaning of educational change*. New York: Teachers College Press.

Fullan, M. (2001b). *Leading in a culture of change*. San Francisco, CA: Jossey-Bass.

Fullan, M. (2003). *The moral imperative of school leadership*. Thousand Oaks, CA: Corwin Press.

Glickman, C., Gordon, S., & Ross-Gordon, J. (2001). *Supervision and instructional leadership: A developmental approach*. Needham Heights, MA: Allyn and Bacon.

Guskey, T. R. (2000). *Evaluating professional development*. Thousand Oaks, CA: Corwin Press.

Guskey, T. R. (2003). What makes professional development effective? *Phi Delta Kappan, 84*(5), 748–750.

Hall, P. (1993). Policy paradigms, social learning and the state: The case of economic policymaking in Britain. *Policy Studies Review, 25*(3), 128–152.

Hightower, A. (2002). *San Diego City Schools: Comprehensive reform strategies at work* (CTP Policy Brief No. 5). Seattle, WA: Center for the Study of Teaching and Policy, University of Washington.

Howlett, M., & Ramesh, M. (1995). *Studying public policy: Policy cycles and policy subsystems*. Don Mills, ON: Oxford University Press.

Leithwood, K., Jantzi, D., & Mascall, B. (2002). *A framework for research on large-scale reform.* Paper presented at the Annual Meeting of the American Educational Research Association in New Orleans.

Levin, B. (2001). *Reforming education: From origins to outcomes*. New York: RoutledgeFalmer.

Levin, B., & Wiens, J. (2003). There is another way: A different approach to education reform. *Phi Delta Kappan, 84*(9), 658–664.

Lieberman, A., & Miller, L. (1999). *Teachers – Transforming their world and their work*. New York: Teachers College Press.

Lipsky, M. (1980). *Street-level bureaucracy: Dilemmas of the individual in public services*. New York: Russell Sage Foundations.

Marzano, R. (2006). *Classroom assessment and grading that work*. Alexandria, VA: Association for Supervision and Curriculum Development.

Marzano, R., Pickering, D., & Pollock, J. (2001). *Classroom instruction that works: Research-based strategies for increasing student achievement*. Alexandria, VA: Association for Supervision and Curriculum Development.

O'Connor, K. (2002). *How to grade for learning: Linking grades to standards*. Thousand Oaks, CA: Corwin Press.

O'Connor, K. (2007). *A repair kit for grading: 15 fixes for broken grades*. Portland, OR: Educational Testing Service.

Organization for Economic Co-operation and Development. (1994). *Quality in teaching*. Paris: OECD Centre for Educational Research and Innovation.

Reeves, D. (2007). From the bell curve to the mountain: A new vision for achievement, assessment and equity. In D. Reeves (Ed.), *Ahead of the curve* (pp. 1–12). Bloomington, IA: Solution Tree.

Richardson, V. (2003). The dilemmas of professional development. *Phi Delta Kappan, 84*(5), 401–407.

Rose, R. (1993). *Lesson-drawing in public policy*. Chatham, NJ: Chatham House Publishers.

Sanders, J. (1997). Applying the personnel evaluation standards to teacher evaluation. In J. Stronge (Ed.), *Evaluating teaching: A guide to current thinking and best practice* (pp. 91–104). Thousand Oaks, CA: Corwin Press.

Scherer, M. (2001). How and why standards can improve student achievement: A conversation with Robert J Marzano. *Educational Leadership, 59*(1), 1–7.

Schon, D. (1987). *Educating the reflective practitioner*. San Francisco, CA: Jossey-Bass.

Sparks, D. (2002). *Designing powerful professional development for teachers and principals*. Oxford, OH: National Staff Development Council.

Stiggins, R., & Chappuis, J. (2005). *Classroom assessment for student learning: Doing it right—Using it well*. Portland, OR: Educational Testing Service.

Stigler, J., & Hiebert, J. (1999). *The teaching gap: Best ideas from the world's teachers for improving education in the classroom*. New York: The Free Press.

Stronge, J. (2002). *Qualities of effective teachers*. Alexandria, VA: Association for Supervision and Curriculum Development.

Tomlinson, C. (2007). Learning to love assessment. *Educational Leadership, 65*(4), 8–13.

Webber, C.F., Aitken, N., Lupart, J., & Scott, S. (2009, May). *The Alberta student assessment study final report* (Report). Edmonton, AB: Alberta Education.

Wilson, S., Darling-Hammond, L., & Berry, B. (2001). *Connecticut's story: A model state teaching policy* (CTP Policy Brief No.4). Seattle, WA: Center for the Study of Teaching and Policy, University of Washington.