Formative Assessment: A Systematic and Artistic Process of Instruction for Supporting School and Lifelong Learning

Ian Clark

University of Washington

Abstract

Formative assessment is a potentially powerful instructional process because the practice of sharing assessment information that supports learning is embedded into the instructional process by design. If the potential of formative assessment is to be realized, it must transform from a collection of abstract theories and research methodologies and become a creative and systematic classroom practice. Policy-makers and school administrators must support this transition from theory into practice, particularly in the early stages of professional adaptation, and design assessment systems that teachers may internalize and enact efficiently. The article explores the hypothesis that many public school teachers are 'trapped' within environments which deter them from enacting open and inventive social learning strategies in their own classrooms, which when implemented have great potential to support autonomous learning, realize achievement, and create economically productive lifelong learners. This article therefore reviews the literature on formative assessment in practical settings, and investigates the extent to which teachers perform the basic functions of gathering and using evidence to further learning and development in pursuit of the lifelong learning competencies that are essential in the 'new economy.'

CANADIAN JOURNAL OF EDUCATION 35, 2 (2012): 24-40

Résumé

L'évaluation formative est un processus éducatif potentiellement fort car la pratique qui consiste à partager l'information issue de l'évaluation et qui soutient l'apprentissage, est intégrée dans le processus éducatif par définition. Le potentiel de l'évaluation formative peut être atteint si une transformation s'opère à partir d'un ensemble de théories abstraites et de méthodologies de recherche, pour devenir une pratique créative et systématique en classe. Les décideurs politiques et les administrateurs scolaires doivent soutenir cette transition de la théorie à la pratique, en particulier dans les premiers stades de l'adaptation professionnelle, et concevoir des systèmes d'évaluation que les enseignants peuvent internaliser et utiliser de manière efficace. L'article explore l'hypothèse que de nombreux enseignants des écoles publiques sont «piégés» dans des environnements qui les dissuadent d'adopter systématiquement des stratégies d'apprentissage innovantes, constructivistes et sociales dans leur propre salle de classe, qui lorsqu'elles sont mises en pratique, ont un fort potentiel afin de soutenir l'apprentissage autonome, favoriser la réussite scolaire, et créer des apprenants économiquement productifs sur le long terme. Cet article passe donc en revue la documentation existante sur l'évaluation formative dans des situations concrètes, et examine dans quelle mesure les enseignants exercent les fonctions de base telles que la collecte et la mise à profit de données probantes afin d'étendre l'apprentissage et développer sur le long terme des compétences d'apprentissage essentielles dans la «nouvelle économie».

Formative Assessment: A Systematic and Artistic Process of Instruction for Supporting School and Lifelong Learning

Introduction

Formative Assessment: A Quiet Revolution in a Rapidly Evolving Century

> In today's world, adaptability, creativity and innovativeness appear to be preconditions for organizations and individuals to thrive…we need people who aren't focused only on payoffs but do the best they can to learn, adapt, improve… (Benkler, 2011, p. 85)

The theme of lifelong learning, is an issue of global interest paralleled by a flourishing interest in formative assessment. Hutchinson and Hayward (2005) describe this trend as a "quiet revolution," one which has across time brought about the transformation of education by embedding "the theory of formative assessment" (Black & Wiliam, 2009) into the policy frameworks of a number of nations (Organization for Economic Cooperation and Development [OECD], 2005; 2008; Educational Testing Service [ETS], 2009). From the perspective of this paper it is important to understand how schools must work in support of Benkler's vision of a cooperative society through the consistent and effective implementation of formative assessments that support lifelong learning competencies. Bransford, Derry, Berliner, Hammerness, and Beckett (2005) note the difference between 'routine experts' and 'adaptive experts,' implying that for students to become effective learners they need to be taught by teachers who have received initial and continuing training on instructional methods which help them to adapt their teaching to meet the needs of their students (Black & Wiliam, 1998a; 1998b; Vogt & Rogalla, 2009). Routine experts have a range of core competencies which they consistently deploy across the span of their lifetimes to attain increasing efficiency. Adaptive experts however have core competencies which are consonant with Benkler's adaptive, creative person. They are more likely to adapt those core competences and innovate new approaches that better equip them to capitalize on opportunities and solve problems which could otherwise appear mystifying. "Adaptive expertise is discussed as the gold-standard for learning in *How People Learn*" (National Research Council [NRC], 2000 cited in Bransford et al., 2005, p. 49). This is because adaptive experts in classroom-assessment are largely in favor of dimensions of efficiency and innovation**,** processes which are not mutually exclusive but intertwined as the basic features of good teaching. Research suggests that for effective formative assessment to take place, teachers need to develop adaptive expertise—short-cycle adaptation of teaching to meet the needs of the students being the key to formative assessment practice (Black & Wiliam, 1998b). Administrators and policy-makers need to support teachers as they adapt their core practices (such adaptation of practice should result in teachers being more flexible in their approach to instruction and assessment), particularly in the short term when they are most susceptible to decreases in their efficiency.

The State of Art at the Turn of the 21st Century

Why 'Art'?

An examination of the 'art' of instruction is an important theme of this article because the notion of teaching as an art form has direct relevance to the successful enactment of formative

assessments. Macintyre, Buck, and Beckenhauer (2007) take an increasingly popular holistic perspective, arguing that the "necessary interpretive eye and capacity to act" (or react) to evidence of learning during interaction requires "artistic teaching visions, attending to the creation of student meaning on an individual and collective basis" (p. 5). The authors believe that "artistic teaching visions offered glimpses into how formative assessment use holds potential to restore the participatory dynamic integral to learning" (p. 1) by linking the process to the product to the learner. Macintyre et al. (2007) found that despite rigorous in-service training on student reasoning and multiple methods for connecting content and learner, efforts to instill "artistry" into teacher practice "began to transcend these efforts as an impoverished need we saw as critical to addressing the formative assessment hurdle" (p. 5). Macintyre et al. (2007) found that participating teachers lacked the artistic understanding to envision how to link process, product, and learner in order to conceive of a holistic schematic for effective teaching. Analogously, Eisner (2005) has for many years opposed methodological reductionism, instead describing teaching as an artistic process that arises from the qualitative judgments of teachers in a way consonant with Black and Wiliam's (2009) "moments of contingency" and Schön's (1987) "indeterminate zone of practice," which requires teachers *and students* to regulate learning by "thinking on their feet." Eisner parallels the patterns of effective instructional practice to musical composition and painting:

> The medium and sensory modality differ, but the business of composing relationships remains. To succeed the artist needs to see, that is, to experience the qualitative relationships that emerge in his or her work and to make judgments about them. (Eisner, 2005, p. 208)

Macintyre, Buck, and Beckenhauer (2007) believe that the process described by Eisner:

> is integral to scientific inquiry. Inquiry is a creative enterprise, resisting imposed routine, demanding reason alongside ongoing judgments, consideration of alternatives, openness, and inventiveness. Such qualities are integral to formative assessments, prompting and furthering learning. (p. 6)

It is perhaps that the 'flair' required for such creative enterprise is the difference between mere competence and technical artistry.

The Development of Formative Assessment: Research and Practice

In 2001, Dixon and Williams of Auckland University in New Zealand presented a paper at the Annual Conference of the British Educational Research Association. The timing of the paper establishes a useful baseline for where formative assessment was in terms of effective use among practitioners at the turn of the 21st century. Their findings were introduced as follows:

> Formative assessment is not well understood by teachers and is weak in practice. Recent research supports the view that conceptually teachers are confused about the nature, purpose and effect of formative assessment. (Dixon & Williams, 2001, ¶1)

Dixon and Williams (2001) found that teachers attributed great importance to formative assessment and its potential to enhance instruction and develop essential lifelong learning competencies among students. However, when asked to articulate their practice in more detail,

they were not able to explain clearly how they used the assessment information gained to enhance children's learning. If practitioners are to conceptualize formative assessment practices as creative and participatory, they need time to translate theory into practice . This transformation will only occur through programs of "development and dissemination which are matched to the capacity of teachers to take ownership of change, and at the same time to rebuild their theories in a form that supports and gives coherence to practice" (Black, 2000, p. 410). As the special advisor to the Canadian premiere, Michael Fullan (2007) notes in the paper 'Change the Terms for Teacher Learning':

> The notion that external ideas alone will result in changes in the classroom and school is deeply flawed as a theory of action. I am not only referring to irrelevant or poorly conducted professional development, but also to sessions that meet the highest standard of adult learning. These activities are not useless, but they can never be powerful enough, specific enough, or sustained enough to alter the culture of the classroom and school. (p. 35)

The momentum for formative assessment into the 21st century in UK, continental Europe, and Australasia was facilitated by a number of key reports published in the 1990s. For example, the Assessment Reform Group (ARG, 1999) based at Cambridge University present a five-point 'blueprint' for a shift from traditional perspectives on instruction toward a culture which emphasizes inquiry and participation by highlighting five key objectives drawn from both cognition and constructivist theories as follows: i) a recognition of the profound influence instruction and specifically assessment has on the motivation and self-esteem of pupils; ii) adjusting teaching to take account of the results of assessment; iii) the provision of effective feedback to pupils; iv) the active involvement of pupils in their own learning; v) the need for pupils to be able to assess themselves and understand how to improve. Particular emphasis is placed upon the sharing of learning goals with pupils; encouraging equality and mutuality between students; involving pupils in self-assessment; and providing feedback which leads to pupils recognizing their next steps and how to take them. The ARG (1999) complete their recommendations for improved instruction by emphasizing several undesirable tendencies exhibited by classroom practitioners: a) a tendency for teachers to assess quantity of work and presentation rather than the quality of learning; b) a focus on marking and grading at the expense of providing advice for improvement, which tends to lower pupils' self-esteem; c) a strong emphasis on comparing pupils with each other, which demoralizes the less successful learners; d) teachers' feedback to pupils often attempts to serve managerial and social purposes rather than helping them to learn more effectively.

Harlen and James (1997) worked alongside teachers in the UK and found that, where formative assessment was occurring, teachers were unaware of it. Conversely, when teachers believed they were assessing formatively they were in reality completing continuous summative assessment which they then used primarily for reporting (not planning) purposes. Foundational studies by Bell and Cowie (1997); Black (1986, 1993); Harlen and Qualter (1991); and Nitko (1995) reported very similar findings. In her work in New Zealand, Baker (1995)found that teachers did not view assessment as integral to teaching and learning; rather, they saw it as an additional task which bore little relationship to classroom activity. Subsequently, this compelled them to dislike and even disdain classroom assessment. For these teachers, there was a feeling of being overloaded with assessment requirements. The traditional 'instruct and summarize' cycle has meant that "while teachers are usually conscientious about marking student work they often

fail to offer guidance on how work can be improved" (Hallam *et al,* 2004, para. 2.4). In attempting to clarify the links between formative assessment and constructivist theories of learning, the literature consistently emphasizes the *sine qua non* of formative assessment as sharing learning goals with learners (Black, 2000; Black, Harrison, Lee, Marshall & Wiliam, 2003; Frederiksen & Collins, 1989; Mansell, James & the ARG, 2009; Sadler, 1989; Shepard, 2000; Stiggins, 2007). While most teachers discussed the importance of working with and alongside children, in most of those cases their discussion did not include sharing learning goals or the importance of supporting individual learning by providing formative feedback.

Dixon and Williams (2001) reported that in the majority of schools where teachers were interviewed, a significant amount of time had been spent on the development of exemplars and benchmarks in various curriculum areas. Implicit within this endeavour was the belief that the benchmarks themselves could be utilized in a formative way, yet no training took place which focused upon the ways they could guide students to move towards these benchmarks. A second and perhaps more significant issue arose from benchmark and interim assessments being erroneously packaged as formative (Popham, 2006). Wiliam (2004) calls such assessments "early warning summative" and Shepard (2005) remarks that the individual profile data from these assessments are not directly formative because: a) the data available are at too gross a level of generality and b) feedback for improvement is not part of the process. Further, while there is debate about the value of immediate versus delayed feedback, there is consensus in the assessment community that learning benefits are more evident when "test results are available quickly enough to enable teachers to adjust how they're teaching and students to alter how they're trying to learn" (Popham, 2006, p. 86). Indeed, even the most streamlined and therefore timely interim and benchmark assessments are pedestrian when one considers the opportunities to gather evidence of student learning through "spontaneous" and "scaffolded" dialogue (cf. Vygotsky). While a number of teachers questioned the relevance of professional development activities which focused on non-formative assessments, there was a general sense that they lacked the interpretive frameworks which would enable them to innovate effective assessments (Dixon & Williams, 2001). Dixon and Williams (2001) found that those who did articulate their concerns about benchmarking and interim assessments had either graduated more recently or undergone study related to assessment in their pursuit of further qualifications. As a result of their greater understanding of formative assessment they were critical of professionals from outside their school who were helping them in particular with benchmarking. They implied that these professionals lacked knowledge about formative assessment and that this had not assisted the process. As a result, these teachers did not consider benchmarking to be useful in helping children with their learning.

The Need for 21st Century Theoretical Transformation

Transformation and Inertia

Heritage (2007) highlights the inertia of educational reform and argues for a reassessment of the pedagogical principles upon which current perspectives on instruction and assessment are founded:

> after more than a hundred years of exhortations and a significant body of research on the topic, the idea that assessment and teaching are reciprocal activities is still not firmly situated in the practice of educators. Instead, assessment is often viewed as something in

> competition with teaching, rather than as an integral part of teaching and learning. (p. 140)

This observation is confirmed by Volante, Drake, and Beckett (2010), who assert, "Although research has clearly shown that formative assessment can enhance student success, there is firm evidence of a research-practice divide. Too many teachers are failing to utilize the full cadre of formative assessment practices available to them" (p. 44). Volante and colleagues undertook a three-year longitudinal study in elementary and secondary schools within Ontario, finding that teachers continued to over-emphasize summative assessment methods (i.e., tests, quizzes, projects), with only a minority of teachers using formative assessment techniques on a consistent basis. The perennial question is why? "Teachers cited a number of factors as constraints to practice, including a lack of instructional leadership, poor initial teacher training, and resistance from parents and students to more innovative formative assessment strategies" (p. 45).

The turn of the 21st century has seen the national policy frameworks of many nations placed under scrutiny, not least for the development and implementation of formative assessments which are designed to provide information for decision-making at the classroom level (OECD, 2005; 2008; ETS, 2009). At the same time, the research and practitioner communities have begun to rebuff the traditionally accepted values of scientific rationalism or 'scientism' that underpin the frequent summative evaluations reauthorized by the No Child Left Behind Act (2002). Scientism posits a simple relationship: that a good school will achieve good performance outcomes. However, this logic is challenged with a sense of increasing urgency by certain aspects of the research community on three main platforms: a) high-stakes testing creates anxiety and disaffection among students (ARG, 1999; Harlen 2006; Harlen & Deakin Crick, 2003); b) the intense focus on testing de-skills, de-motivates, and de-professionalizes teachers and diminishes the quality of public education; and c) the political emphasis on school-level data narrows the instructional content taught to students, distorting how and what students are taught (Abrams, 2007; Shepard, 2000). Ruthven's (1994) thematic review of fifty presentations at the 2nd International Congress on Mathematical Instruction (ICMI) seminar confirm the views expressed by Shepard (2000), Abrams (2007), and Frederiksen and Collins (1989) who are highly critical of summative tests which "subvert" teaching and lead to an emphasis on the memorization of facts and rote sequences at the expense of reasoning skills.

Despite the clear articulation of these issues, 'scientific rationalism' continues to drive educational policy in many nations. In the U.S., the persistence of a current policy direction founded upon scientism prompted the National Association of State Boards of Education (NASBE, 2009) to remark in a study on assessment systems for the 21st century learner that "a growing majority of testing experts and analysts now believe that education cannot be transformed under the constraints of the current state assessment and accountability systems" (p. 3). Brian Gong (2007), highlights the inefficiencies created by outdated"19th century" policy perspectives which effectively rebuff inquiry as a creative process:

> I can imagine that unless an investment is made, educational testing in 20 years will be as hobbled by a lack of imagination and by 19th century measurement theories as it is today. Reauthorization should look to the future as well as try to make mid-course corrections to the present. (p. 4)

Like Gong, Freedman (1992) suggests that reductionism remains rooted in 19th century scientific methodology. Reductionist methodologies, which seek to reduce the world of physical

phenomena to a finite set of fundamental equations (Dyson, 1992) do not adequately address the study of complex, non-linear, spontaneous, and multi-dimensional social interactions upon which learning 'systems' are based., Macintyre, et al. (2007) emphasize formative assessment as a complex and dynamic process – as such, it is a process tailored for the turbulent, unpredictable event that is the learning moment. Macintyre et al. express a preference for Dewey's (1934) term 'esthetic recurrence,' which describes a rhythm in the learning process, one that combines past events with future possibilities, which blend together to take learning forward. As Macintyre and colleagues approached the end of their two-year study into the use of formative evidence they noted that "esthetic recurrence was what teachers gradually attended to, manifesting relationships…This capacity to perceive relationships among parts was what teachers struggled to interpret and act on throughout the study" (2007, p. 15; cf. Eisner). Dewey's holistic notion of 'esthetic recurrence' is consonant with Eisner's rejection of reductionism and the view of teaching as an art-form. Esthetic recurrence exhibits three key features:

- It looks to the individuality present in each learning artifact seeking personal learning connections.
- It reaches out to the relations, associations, and interactions with other individuals, expanding the whole.
- It presents itself in different contexts and with differing learning consequences so that every recurrence was "novel as well as a reminder" (Dewey, 1934, p. 169).

Esthetic recurrence therefore "looks to the individuality present in each learning artifact seeking personal learning connections but also reaching out to the relations, associations, and interactions with other individuals" (cf. Benkler, 2011; Macintyre, Buck & Beckenhauer, 2007, p. 15). It is perhaps inevitable that when teachers are trained to work within an environment which advances scientific rationalism as the prevailing educational philosophy they find it a challenge to "compose" qualitative learning relationships (Eisner, 2005) which capture the significance of past, present and, future learning progressions for each student

Recent Research into Classroom Practice

Many teachers agree in principle on the potential of formative assessment to improve instruction. However, a significant majority lack the level of declarative knowledge required for the development of procedural efficiency (Black & Wiliam, 1998b; Dixon & Williams, 2001; Herman, Osmundson & Silver, 2010). As such assessment practices "remain at the level of intuitive practice, a product of the teacher's personal relations and style" (Pryor & Torrance, 1996, p. 217). One consequence of this lack of information fluidity is the inconsistent adoption and application of formative assessments. Michael Fullan (2007b) suggests a general cause for the lack of development in formative assessment and other potentially beneficial instructional practices is "that schools and the individuals in them may lack the knowledge, skills and personnel to implement any change agenda" (p. 58). Black & Wiliam (2006) emphasize the scale of the task facing schools and teachers, calling it a "formidable" task and note the improbability of transformation without the committed investment of resources and expertise from external sources.

Such findings are consistently reiterated in more recent field research into the formative use of evidence. For example, Campos and O'Hern (2007) implemented a study which engaged first and fifth grade teachers and their students in the formative assessment process in order to

assess the post-intervention impact. Teachers had consistently mentioned students' lack of declarative knowledge and procedural skills to complete classroom assignments which resulted in disappointing test scores. Prior to intervention, the analysis of probable cause data revealed a number of factors which diminished the effectiveness of Mathematics instruction: i) students were unclear about what they needed to know; ii) systems beliefs about testing focus on sorting students rather than motivating them; iii) students were not taking responsibility for their learning, and iv) students did not have adequate, descriptive feedback on their work and were not encouraged to reflect on their learning. The pre-intervention analysis suggested strongly that teachers and students would benefit from the process of formative assessment. Accordingly, the post-intervention data indicated an increase in student involvement in their learning in the area of mathematics. Many students were more aware of grade level mathematics goals. They were able to talk about the mathematics goals and monitor their progress toward these goals. Students also displayed the important strategy of soliciting help when needed in class (Campos & O'Hern, 2007, p. i; Pintrich, 1999; 2004; Vygtosky, 1978; Zimmerman & Pons, 1986). Social assistance is crucial to academic success. For example, Zimmerman and Pons (1986) found that of the high-achievers interviewed, 50% asked for feedback from peers and 35% from adults (teachers and parents). In contrast, the key lifelong learning competence of social engagement was rarely pursued among low-achievers with only 23% seeking assistance from peers and 8% from adults. Zimmerman and Pons (1986) implicitly support the use of formative assessments by their concluding remark: "The present results suggest that theoretical conceptions of students as initiators, planners, and observers of their own instructional experiences have empirical and practical merit" (p. 626). Campos and O'Hern (2007) found that participation increased students' awareness, enabling learners to become self-regulatory and autonomous agents. A most important observation related to the development of self-regulated learning (SRL) strategies — the over-arching goal of formative assessment with the potential to support learning, improve outcomes, and instill lifelong learning competences (Black & Wiliam, 2009; Bose & Rengel; 2009; Clark, 2012; Irving, 2007; Nicol & Macfarlane-Dick, 2006). It was acknowledged that the intervention yielded positive results, in no small part to the enthusiasm of the school principal, who was described as "an invaluable resource" (p. 10).

Collection and Use of Evidence

However, recent research interventions have not always been successful and serve only to highlight the extent to which weak assessment is embedded in the day-to-day work of many teachers. For example, in a recent study, Macintyre et al. (2007) focused on the "lived experience" of teachers and emphasized open and spontaneous dialogue as "the fundamental means utilized to initiate and extend understandings of inquiry in a continuous responsive interchange" (p. 3). All participating teachers taught at the sixth grade level, followed the same district-defined curriculum, and their students completed the same district-wide Science exams. School district objectives for sixth grade science content formed the shared purpose. The specific aim of the study "was to make the inquiry process discernible to teachers fostering growth in learner understandings and teacher insights into these understandings" (p. 4). A pre-intervention analysis discovered that "only 5% of the formative assessment artifacts built into the process of inquiry learning were used by teachers to adapt instruction so that student thinking was furthered" (p. 5). The finding was described as remarkable by the researchers, who continued, "despite the large number of artifacts collected, they actually provided little insight into their role in forming and reforming learning and in fact were seen by teachers as quite disconnected from the actuality of the student inquiries conducted" (p. 5). The study spanned two years (Phase 1

and Phase 2). In year one, 4000 artifacts were examined, and in year two, 3000. The *usefulness* of the artifacts as materials which located the intersections of the learning situation with students' understandings was illustrated by a focus on two Phase 1 classrooms.

> One classroom contained 101 student-generated artifacts. Of these, 27% were coded as artifacts that provided no understanding of student learning, 49% were coded as artifacts that provided little understanding, 25% were coded as artifacts that revealed a limited understanding of students' knowledge construction, and none of the artifacts provided a thorough understanding of student learning. (Macintyre et al., 2007, p. 13)

In another classroom, 378 student documents were collected.

> Of these, 70% were coded as artifacts that provided no understanding of student learning, 3% were coded as artifacts that provided little understanding, 25% were coded as artifacts that revealed a limited understanding of students' knowledge construction, and 2% of the artifacts provided a thorough understanding of student learning. (Macintyre et al., 2007, p. 13)

Unfortunately, the quantity of evidence that provides either a limited, or preferably, a thorough understanding of student learning was found to be inadequate. The absence of quality assessment data emerged as a generalizable theme among the teachers in the study: "the percentages for other teachers reflected similar percentages of the latter two categories, even if the percentage of artifacts was significantly different for the first two categories" (p. 13). These findings prompted the researchers to remark that they were "haunted" by Sidorkin's (2001) "characterization of education as the production of useless things ... as we attended to these artifacts in relation to the 2-year research process" (p. 15). Evidence collection was undertaken in what teachers believed to be a satisfactory manner, and indeed many were proud of their formative interventions. However, the researchers could not agree, assessing such enactments as superficial in that they did not further the learning process. In a U.S study (Herman, Osmundson & Silver, 2010), elementary science teachers (n=40) indicated that they engaged in a wide variety of assessment. For example, the provision of individual, written feedback, using scoring rubrics, recording observations, checking student understandings at the end of investigations, and using the data to guide subsequent weekly instruction. However, the researchers noted a discrepancy between teachers' opinion of themselves as competent assessors and the observed use of evidence to further student learning:

> Rather, the general lack of relationship between the quality of teachers' based on their ability to analyze student understanding and respond with instructional next steps, and teachers' use of the assessment process, based on self report survey data … suggest that it may be important to differentiate teachers' engagement in the process of assessment from the validity of the inferences they are able to draw and use from that process. (p. 17)

In their report, Herman and colleagues suggest that "one interpretation of study findings is that teachers 'talk the talk,' but findings on the teacher knowledge and the quality of their assessments suggest that they need help to more fully 'walk the walk'" (p. 18).

As noted earlier, the cornerstone of formative assessment is the discussion and development of assessment criteria *among* students. In the U.K., Tiknaz and Sutton (2006) found

that peer discussions were extremely rare (once or twice a year). This was essentially related to three factors. The first major barrier is the difficulty of finding time for peer-assessment when there was a perceived need to meet the statutory requirements of the Programme of Study. A second major barrier to the sharing of assessment criteria is the difficulty that some teachers experience in translating assessment criteria into a language that their pupils can understand. The third barrier is teachers' preconceptions about the ability of their students to engage in the assessment process. Teachers appeared to be reluctant to pass the assessor roles to students, in particular to Year 7 groups (11 year olds), whom they thought did not have the maturity to conduct an objective self- and peer-assessment. Student autonomy, which refers in this case to the extent to which pupils assumed responsibility to set and review their own targets, was permitted by only three of the 12 teachers. Participating teachers were mostly skeptical about students' ability to assess each other and set targets. As a consequence, student engagement in thinking about their own learning was very much teacher-directed (Tiknaz & Sutton, 2006). Of particular interest is the finding that ten teachers perceived that lower 'ability' students tended to set unrealistic targets or those that were not meaningful. Advocates of formative assessment claim that there are two fundamental problems with this perspective: i) ability is not a valid issue in a formative assessment classroom. Instead it is the belief that one's conscious efforts will result in success, i.e., self-efficacy (cf. Dweck; Bandura); and ii) an important early claim (ARG 1999; Black & Wiliam, 1998a; 1998b) supporting the use of formative assessment specifically noted that it is of most benefit to low-achieving students as it fosters much needed meta-cognitive skills required for the development of lifelong learning competencies.

The whole point of collecting evidence of learning is to then use it diagnostically to ascertain students' existing knowledge and then plan next steps for individual learning progressions. In the U.S., Herman et al. (2010) directly addressed this key issue in CRESST Report 770. Teachers were assessed for their skill at prescribing next instructional steps for the student(s), the indicator of quality of use. Teachers were rated on a scale of 0-21 and performance varied greatly. At the low end of the spectrum, four teachers scored a total of only two points; at the other end of the continuum, the three highest scoring teachers scored 13. The average score was 6.7 (SD = 3.2), with a range of 2 to 13 points. The scores of the great majority, 80% of the teachers, were clustered around score points "1" and "2" suggesting that most teachers tended to rely on general approaches to subsequent instructional planning. Herman and colleagues's findings are consistent with other recent studies on teacher pedagogical content knowledge (e.g., Heller, Daehler, Shinohara & Kaskowitz, 2004; Heritage & Vendlinski, 2006).

Cornerstone

Training and Preparedness

The importance of training methods pursuant of creative, responsive and systematic instruction has been mentioned earlier in the article (Black, 2000; Fullan, 2007a; Vogt & Rogalla, 2009). Therefore, in concluding, it seems logical to review the hypothesis that the inconsistent adoption of formative assessment may be resolved by teacher training methods which prepare teachers for its use. Black et al. (2003) found that teachers who received training in formative assessment and were provided with ongoing, collaborative support were able to make substantial changes within their classrooms. Macintyre et al. (2007) recommend that teachers are trained so they develop insights in four key areas: seeing, relational knowing, mindful embodiment, and continual assessment.

i) Working with *seeing* positioned teachers to foster inner attention in students so that the control for learning came from within the learning situation itself;
ii) Working with *relational knowing* positioned teachers to find intersections among student, teacher, and subject matter;
iii) Working with *mindful embodiment* positioned teachers to concretely feel the lived terms of inquiry alongside their students;
iv) Working with *continual assessment* positioned teachers to support and enhance learning processes.

This framework promotes practices that teachers may use to collect evidence and plan for the next steps in the learning progression. However, in their Canadian study Volante et al. (2010) found that few

> faculties teach courses on assessment and evaluation; rather, assessment is embedded into teachable subject areas. The trouble with this design is that not all faculty members have expertise in assessment and evaluation, and therefore the assessment content is not properly infused and is often neglected in teachable subject courses. (p. 46)

It is concerning that educators continue to hold assessment perspectives which fail to promote learning as a creative enterprise characterized by the consideration of alternatives, openness, and inventiveness (Macintyre et al., 2007). Stiggins (2004) refers to the U.S. when he remarks, "in districts, schools, and classrooms across the nation, educators still assess student learning the way their predecessors did 60 years ago because they have not been given the opportunity to learn about new insights and practices" (p. 22). This raises questions about the efficacy of staff development opportunities and brings into question the ability of those professionals who are leading in-service development. Gipps (1994) and Grossman et al. (2008) concur, emphasizing that staff undertaking leadership roles in the development of school-wide assessment practices need to have expert knowledge of current theories of classroom assessment, how these relate to the ways in which children learn including their impact upon lifelong learning competencies.

Parr and Timperley (2008) note that "despite the current notion of assessment as integral to school and class programs, there are questions about the preparedness of schools and teachers to engage in the process" (p. 58). Where teachers are found to be unprepared to conduct formative assessment it may be that the routine use of student achievement data in relation to decision-making may not be a valued part of the professional culture at that school (Robinson & Sinnema 2006; Sinnema 2005). In circumstances where schools do place value to formative assessments it raises the question of whether typical teachers possess the requisite creative skills and artistic vision to work with data systematically collected with respect to student performance (Black & Wiliam, 2006; Black & Wiliam, 2009; Macintyre et al., 2007; Parr & Timperley, 2008; Parr, Timperley, Reddish, Jesson, & Adams, 2006). Volante et al. (2010) emphasize that "instructional leadership in assessment is the key to successful data decision making. As one administrator from our study said, **...** extracting next steps from assessment data is the cornerstone to effective leadership, instruction and student growth" (p. 47).

References

Abrams, L. M. (2007). Implications of high-stakes testing for the use of formative classroom assessment. In H. McMillan (Ed.), *Formative classroom assessment: Theory into practice* (pp. 8–28). New York, NY: Teachers College Press.

Assessment Reform Group (ARG). (1999). *Assessment for learning: Beyond the black box.* Cambridge, UK: Cambridge University, School of Education.

Ayala, C. C. (2005). Formative assessment guideposts. *Science Scope, 28*(4), 26-28.

Baker, F. (1995). Moderation. *Teaching: The Magazine for Professionals in Education*, *33*(1 & 3).

Bell, B., & Cowie, B. (1997). *Formative assessment and science education*. Report of the Learning in Science Project (Assessment). Centre for Science, Mathematics, Technology Education Research. University of Waikato.

Benkler, Y. (2011). The unselfish gene. *Harvard Business Review*, *89*(7/8), 77 – 85.

Black, H. (1986). Assessment for learning. In D. Nuttall (Ed.), *Assessing educational achievement* (pp. 7-18). Lewes: Falmer Press.

Black, P. (1993). Formative and summative assessment by teachers. *Studies in Science Education*, *21*, 49-97.

Black, P. (2000). Research and the development of educational assessment. *Oxford Review of Education*, *26* (3&4), 407-419.

Black, P., Harrison, C., Lee, C., Marshall, B., & Wiliam, D. (2003). *Assessment for learning: Putting it into practice.* Buckingham, UK: Open University Press.

Black, P., & Wiliam, D. (1998a). Assessment and classroom learning. *Assessment in Education*, *5*(1), 7–73.

Black, P., & Wiliam, D. (1998b). *Inside the black box: Raising standards through classroom assessment.* London: School of Education King's College.

Black, P., & Wiliam, D. (2006). Assessment for learning in the classroom. In J. Gardner (Ed.), *Assessment and learning* (pp. 9 -25). London: Sage.

Black, P., & Wiliam, D. (2009). Developing the theory of formative assessment. *Educational Assessment, Evaluation and Accountability, 21*(1), 5-31.

Bose, J., & Rengel, Z. (2009). A model formative assessment strategy to promote student-centered self-regulated learning in higher education. *US-China Education Review, 6*(12), 29-35.

Bransford, J., Derry, S., Berliner, D., Hammerness, K., & Beckett, K.L. (2005). Theories of learning and their roles in teaching. In L. Darling-Hammond & J. Bransford (Eds.). *Preparing Teachers for a Changing World: What Teachers Should Learn and Be Able to Do* (pp. 40-87). San Francisco, CA: Jossey-Bass.

Campos, J., & O'Hern, J. (2007). *How does using formative assessment empower students in their learning?* Unpublished Master of Arts in Teaching and Leadership Saint Xavier University & Pearson Achievement Solutions. Retrieved from http://www.eric.ed.gov/PDFS/ED496594.pdf

Clark, I. (2012). Formative assessment: Assessment is for self-regulated learning. *Educational Psychology Review, 24*(2), 205-249.

Crooks, T. (1988). The impact of classroom evaluation practices on students. *Review of Educational Research, 58*(4), 438-481.

Dewey, J. (1934). *Art as experience*. New York: Capricorn.

Dixon, H., & Williams, R. (2001, September). *Teachers' understandings of formative assessment*. Paper presented at the Annual Conference of the British Educational Research Association. Retrieved from http://www.leeds.ac.uk/educol/documents/00002533.htm

Dyson, F. (1995). The scientist as rebel. *The New York Review*, *42*, 31-33.

Educational Testing Service (ETS). (2009). Research rational for *Keeping Learning on Track*® program. Retrieved from http://www.assessmentinst.com/wp-content/uploads/2009/08/11518-klt-pearbro_bw.pdf

Eisner, E. (2005). What can education learn from the arts about the practice of education? In E. Eisner (Ed.), *Reimagining schools: The selected works of Elliot E. Eisner* (pp. 205 – 214). NY: Routledge.

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

Freedman, D. H. (1992). Is management still a science? *Harvard Business Review*, (Nov-Dec), 26-38.

Fullan, M. (2007a). Change the terms for teacher learning. *National Staff Development Council, 28*(3), 35 – 36. Retrieved from http://www.michaelfullan.ca/Articles_07/07_term.pdf

Fullan, M. (2007b). *The new meaning of educational change* (4th ed.). New York: Teachers College Press.

Gipps, C. (1994). *Beyond testing: Towards a theory of educational assessment.* London: Falmer Press.

Gong, B. (2007). *Testimony of Brian Gong*, Executive Director, The National Center for the Improvement of Educational Assessment before the Committee on Education and Labor hearing on the "Miller-McKeon Discussion Draft of the Reauthorization of the Elementary and Secondary Education Act." Washington, D.C., September 10, 2007. Retrieved from http://edlabor.house.gov/testimony/091007BrianGongTestimony.pdf

Grossman, P., Hammerness, K., & McDonald M. (2008). Redefining teaching, re, imagining teacher education. *Teachers and Teaching: Theory and Practice, 15*(2), 273-289.

Hallam, S., Kiston, A., Peffers, J., Robertson, P., & Stobart, G. (2004). *Evaluation of Project 1 of the AiFL Development Programme: Support for professional practice in formative assessment*.

Institute of Education (IoE), University of London. Retrieved from www.scotland.gov.uk/Publications/2004/10/19947/42988

Harlen, W. (2006). The role of assessment in developing motivation for learning. In J. Gardner (Ed.), *Assessment and learning* (pp.61 - 80). London: Sage.

Harlen, W., & Deakin Crick, R. (2003). Testing and motivation for learning. *Assessment in Education*. *10*(2), 169-207.

Harlen, W., & James, M. (1997). Assessment and learning: Differences and relationships between formative and summative assessment. *Assessment in Education: Principles, Policy & Practice,* 4(3), 365-379.

Harlen, W., & Qualter, A. (1991). Issues in SAT development and the practice of teacher assessment. *Cambridge Journal of Education, 21*(2), 141-152.

Heller, J. I., Daehler, K. R., Shinohara, M., & Kaskowitz, S. R. (2004). *Fostering pedagogical content knowledge about electric circuits through case-based professional development.* Paper presented at the annual meeting of the National Association for Research in Science Teaching (NARST), Vancouver, Canada.

Heritage, M. (2007). Formative assessment: What do teachers need to know and do? *Phi Delta Kappan, 89*(2), 140-145.

Heritage, M., & Vendlinski, T. (2006). *Measuring teachers' mathematical knowledge* (CSE Tech. Rep. No. 696). Los Angeles, CA: University of California, National Center for Research on Evaluation, Standards, and Student Testing (CRESST).

Herman, J. L., Osmundson, E., & Silver, D. (2010). *Capturing quality in Formative Assessment practice: Measurement challenges*. CRESST Report 770. National Center for Research on Evaluation, Standards, and Student Testing. Retrieved from http://www.cse.ucla.edu/products/reports/R770.pdf

Hutchinson, C., & Hayward, L. (2005). The journey so far: Assessment for learning in Scotland. *The Curriculum Journal, 16*(2), 225-248.

Irving, K. (2007). *Teaching science in the 21st century: Formative assessment improves student learning*. National Science Teachers Association (NSTA). Retrieved from http://www.nsta.org/publications/news/story.aspx?id=53559

Mansell, W., James, M., & the Assessment Reform Group. (2009). *Assessment in schools. Fit for purpose?* A commentary by the Teaching and Learning Research Programme.

Macintyre, L. M., Buck, G., & Beckenhauer, A. (2007). Formative assessment requires artistic vision. *International Journal of Education & the Arts, 8*(4), 1 – 23. Retrieved from http://ijea.asu.edu/v8n4/

Mercer, N. (1995). *The guided construction of knowledge*. Clevedon: Multilingual Matters.

Mercer, N. (2004). Sociocultural discourse analysis: Analysing classroom talk as a social mode of thinking. *Journal of Applied Linguistics, 1*(2), 137-168.

Morris, P. (1999). *The project on feedback and assessment*: *Final report*. Hong Kong: Hong Kong University.

National Association of State Boards of Education (NASBE). (2009). *A call for balancedsystems of assessment and accountability*. NASBE Study Group on Assessment Systems for the 21st Century Learner. Retrieved from http://www.k12.wa.us/assessment/ClassroomAssessmentIntegration/pubdocs/NASBEAssessmentReport_ExecSumry.pdf

Nicol, D. J., & Macfarlane-Dick, D. (2006). Formative assessment and self-regulated learning: A model and seven principles of good feedback practice. *Studies in Higher Education, 31*(2), 199-218.

Nichols, P.D., Meyers, J. L., & Burling, K. S. (2009). A framework for evaluating and planning assessments intended to improve student achievement. *Educational Measurement: Issues and Practice, 28*(3), 14-23.

Nitko, A. (1995). Curriculum based continuous assessment: A framework for concepts, procedures and policy. *Assessment in Education: Principles, Policy and Practice, 2*, 321-337.

Noddings, N. (2004). High stakes testing: Why? *Theory and Research in Education*, *2*(3), 263-269.

Organisation for Economic Co-operation and Development (OECD). (2005). *Formative Assessment: Improving learning in secondary classrooms*. Centre for Educational Innovation and Research. Paris: OECD Pub.

OECD/CERI. (2008). *Assessment for learning: Formative assessment*. International Conference, Learning in the 21st Century: Research, Innovation and Policy. Paris: OECD. Retrieved from http://www.oecd.org/dataoecd/19/31/40600533.pdf

Parr, J.M., Timperley, H.S., Reddish, P., Jesson, R. & Adams, R. (2006). *Literacy professional development project: Identifying effective teaching and professional development practices for enhanced student learning.* Wellington, New Zealand: Ministry of Education.

Parr, J. M., & Timperley, H. S. (2008). Teachers, schools and using evidence: Considerations of preparedness. Assessment in Education: Principles, Policy & Practice*, 15*(1), 57-71.

Pintrich, P. (1999). The role of motivation in promoting and sustaining self-regulated learning. *International Journal of Educational Research, 31*(6), 459-470.

Pintrich, P. (2004). A conceptual framework for assessing motivation and self–regulated learning in college students. *Educational Psychology Review, 16*(4), 385–407.

Popham, W. J. (2006). Phony formative assessments: Buyer beware! *Educational Leadership, 64*(3), 86–87.

Pryor, J., & Torrance, H. (1996). Teacher-pupil interaction in formative assessment: assessing the work or protecting the child? *The Curriculum Journal, 7*(2), 205-226.

Rex, L. A., & Schiller, L. (2009). *Using discourse analysis to improve classroom interaction*. NY & Abingdon (UK): Taylor and Francis.

Robinson, V.M.J., & C. Sinnema. 2006. *The leadership of teaching and learning: Implications for teacher evaluation.* Paper presented to Annual Meeting of the American Educational Research Association, April 7–11, in San Francisco, USA.

Sadler, D. R. (1989). Formative assessment and the design of instructional systems. *Instructional Science, 18*(2), 119-144.

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

Shepard, L. (2000). The role of assessment in a learning culture. *Educational Researcher, 29*(7), 4–14.

Shepard, L.A. (2005). *Formative assessment: Caveat emptor*. ETS Invitational Conference 2005. The Future of Assessment: Shaping Teaching and Learning, New York. Retrieved from http://www.cpre.org/ccii/images/stories/ccii_pdfs/shepard formative assessment caveat emptor.pdf

Sidorkin, A. (2001). The labor of learning. *Educational Theory, 51*(1), 91-107.

Sinnema, C. 2005. *Teacher appraisal: Missed opportunities for learning*. Unpublished PhD thesis, The University of Auckland.

Sommers, M.K. (2007). *Statement of Mary Kay Sommers*, Ph.D., President, National Association of Elementary School Principals (NAESP), "Reaction to the Miller- McKeon Discussion Draft of the Reauthorization of the Elementary and Secondary Education Act (ESEA)" Washington, D.C. September 10, 2007. Retrieved from http://edlabor.house.gov/testimony/091007MaryKaySommersTestimony.pdf

Stiggins, R. (2004). New assessment beliefs for a new school mission. *Phi Delta Kappan, 86*(1), pp.22-27.

Stiggins, R. (2007). Conquering the formative assessment frontier. In H. McMillan (Ed.), *Formative classroom assessment: Theory into practice* (pp. 8–28). New York, NY: Teachers College Press.

Sullivan, H. S. (1953). *The interpersonal theory of psychiatry.* New York, NY: Norton.

Tiknaz, Y., & Sutton, A. (2006). Exploring the role of assessment tasks to promote formative assessment in Key Stage 3 Geography: evidence from twelve teachers. *Assessment in Education 13*(3), 327–343.

Vogt, F., & Rogalla, M. (2009). Developing adaptive teaching competency through coaching. *Teaching and Teacher Education, 25*(8), 1015-1060.

Volante, L., Drake, S., & Beckett, D. (2010). Formative assessment: Bridging the research-practice divide. *Education Canada, 50*(3), 44 – 47.

Vygotsky, L. S. (1978). *Mind in society: The development of higher psychological processes*. Cambridge, MA, Harvard University Press.

Wiliam, D. (2004). *Assessment and the regulation of learning*. Paper presented at Invited Symposium 'What does it mean for classroom assessment to be valid? Reliable?' at the annual meeting of the National Council on Measurement in Education, April 2004, San Diego, CA. Retrieved from http://www.michigan.gov/documents/mde/WiliamAssessmentarticle_201263_7.pdf

Zimmerman, B.J., & Pons, M.M. (1986). Development of a structured interview for assessing student use of self-regulated learning strategies. *American Educational Research Journal, 23*(4), 614-628.