Formative assessment and the contemporary classroom: Synergies and tensions between research and practice

Louis Volante
Brock University

Danielle Beckett
Brock University

Author's Note

Abstract

Twenty teachers working in elementary and secondary schools were interviewed from 2 school districts in southern Ontario, Canada about their understanding and use of particular formative assessment strategies. Analysis of the interviews followed a constant comparison method and revealed a variety of emerging themes. Results suggested an imbalance in the use of formative assessment methods associated with improvements in student learning and achievement. Many teachers noted tensions in using particular formative assessment strategies such as peer assessment and self-assessment. The discussion focuses on the implications for teacher education reform and in-service professional development so that greater synergy between formative assessment research and practice can be obtained in contemporary classrooms.

Descriptors: Formative assessment; teacher knowledge; professional development.

Résumé

Vingt enseignants travaillant dans des écoles primaires et secondaires, de 2 districts scolaires dans le sud de l'Ontario au Canada, ont été interrogés au sujet de leur compréhension et de leur utilisation de stratégies particulières d'évaluation formative. L'analyse des entrevues a suivi une méthode de comparaison constante et a révélé une variété de thèmes émergents. Les résultats suggèrent un déséquilibre dans l'utilisation de méthodes d'évaluation formative associé à l'amélioration de l'apprentissage des élèves et à leur réussite. De nombreux enseignants ont remarqué des tensions en utilisant notamment des stratégies d'évaluation formative telles que l'évaluation par les pairs et l'auto-évaluation. Le débat se concentre sur les implications d'une réforme de la formation des enseignants et du développement professionnel des enseignants en service, afin qu'une plus grande synergie entre la recherche et la pratique de l'évaluation formative puisse être obtenue dans les salles de classe contemporaine.

Descripteurs : évaluation formative, connaissances des enseignants, développement professionnel.

**Formative assessment and the contemporary classroom:
Synergies and tensions between research and practice**

In recent years, research has highlighted the importance of different phases of assessment and a divergence from the "teach, test, and hope for the best" model that has dominated schools (Earl, 2003; Harlen, 2007; Stiggins, 2008; Volante, 2010). The rationale for this shift has been coupled with many hopeful signs that improvements in classroom assessment will contribute to the improvement of student learning (Black, Harrison, Lee, Marshall, & Wiliam, 2004). Nevertheless, it is only the consistent use of formative assessment (also known as assessment *for* learning) that has shown promise in improving student learning and achievement (Black & Wiliam, 2006; Earl & Katz, 2006; Assessment Reform Group, 2002). Formative assessment practices are ongoing and take place during a lesson or unit of study. Examples might include a student completing a journal reflection, self-assessment of a performance, or submission for a draft of a final assignment. Conversely, summative assessment strategies (also known as assessment *of* learning) are those that primarily serve an evaluative function at the end of a unit or term. Summative assessment methods are typically traditional paper-and-pencil measures such as quizzes, tests, exams, essays, or projects that form a portion of a student's final grade. For example, many secondary students in North America complete a final exam that is worth a significant portion of their final grade. These final exams are used to determine the degree of achievement of specific competencies in particular subject areas such as science, mathematics, geography, history, or English.

Although both forms of assessment serve specific and separate functions, summative and formative assessments are not mutually exclusive in practice. That is, it is the purpose of the assessment, rather than the task, that delineates the form of the assessment (Earl & Katz, 2006). Traditionally, teachers have had a difficult time incorporating various types of assessment in a multi-synergistic and purposeful fashion (Earl & Katz, 2006; Volante, 2010; Wilson, 2004).

Research has suggested that specific formative assessment practices have a direct impact on student learning and achievement. In particular, four large reviews on the impact of formative assessment (Black & Wiliam, 1998; Crooks, 1988; Kluger & DeNisi, 1996; Natriello, 1987) have supported the claim that the use of formative strategies such as questioning techniques, feedback without grades, self-assessment, peer assessment, and formative use of summative assessments can double the speed of student learning (Wiliam, 2007). Even more importantly, formative assessment reduces the achievement gap by helping low achievers the most (Black et al., 2004; Black & Wiliam, 1998). Unfortunately a constricted range of assessment practices, particularly those that emphasize traditional paper-and-pencil summative measures, are being overemphasized within contemporary schools (Earl, 2003; Popham, 2005; Stiggins, 2008; Volante, 2010). Thus, the reform of schools and classroom assessment strategies are intimately connected and the ability to promote diverse formative assessment strategies is paramount to school success (Harlen, 2005; Popham, 2005; Stiggins, 2008; Wilson, 2008).

In order to expand the current research on formative assessment practice, a group of elementary and secondary teachers were interviewed about their self-perceived skill in formative assessment. The interview protocol contained a range of questions that focused on expertise and utilization with various formative assessment methods that are

associated with improvements in student learning and achievement. The primary objective was to identify which practices may be under-used, and more importantly, the key factors that might account for a potential research–practice gap. The results have the potential to inform teacher education reform, in-service professional development, and capacity building efforts geared at transforming classroom practice.

Studying Perceptions

Studying teachers' perspectives of assessment is important because evidence suggests that teachers' perceptions of teaching and learning strongly influence how they teach and what students learn and achieve (Brown, 2004). To illustrate, Kahn's (2000) case study of assessment in secondary school English classes revealed an eclectic array of conflicting assessment practices, seemingly because the teachers held differing perceptions of teaching and student learning. Similarly, research suggests that changes in formative assessment practices are directly related to changes in teachers' attitudes (Dekker & Feijs, 2005). Therefore, it is imperative that researchers and teacher development providers gauge teachers' assessment perceptions before implementing teacher education reforms or professional development programs targeted at in-service teachers.

Theoretical Framework

The theoretical framework for this study was guided by the work of Black Harrison, Lee, Marshall, & Wiliam (2004). This research was used to develop instrumentation that could tap specific formative assessment strategies (questioning techniques, feedback without grades, self-assessment, peer assessment, and the formative use of summative assessments) that are associated with improvements in student learning and achievement. Overall, the researchers used assessment for learning as an overarching construct (Assessment Reform Group, 2002). As well, the assessment for learning construct provides a method for generating specific recommendations that will be useful for policy-makers, district and school staff involved in capacity building initiatives, and teacher educators. Indeed, the relative importance of formative assessment has been recognized as an urgent priority by educational researchers, assessment specialists, and practitioners around the world (Brown, 2004; Dekker & Feijs, 2005; Stiggins, 2002).

Context

One of the major changes in assessment over the last decade is rooted in the use of high-stakes, large-scale testing programs to provide accountability evidence that schools and teachers are meeting public expectations (Decker & Bolt, 2008). As a result, teachers in many western educational jurisdictions (i.e., England, United States, Australia, and Canada) feel pressured by outside sources expecting students to perform well on these summative assessment measures. While the targets and results of national, provincial, and regional agendas may vary, few educators today would dispute the reality that constant pressure imposed on schools to boost students' scores on accountability measures has an influence on classroom instruction. Teaching to the test, narrowing of

the curriculum, neglect of higher-order and critical thinking skills, are some of the frequently cited unintended consequences of an excessive focus on external testing (Popham, 2009; Volante, 2008).

Educational assessment in Ontario more or less falls into two categories: large-scale assessment and classroom assessment. The Educational Quality and Accountability Office (EQAO) is responsible for the creation and administration of large-scale assessments in Ontario while classroom assessment is ultimately at the responsibility of teachers with support from school administrators. Unlike some jurisdictions in the Western world—such as those in select parts of the United States, the United Kingdom, and Australia—there is no formal requirement to use classroom assessment data (also referred to as curriculum-embedded assessment) for accountability purposes in Ontario (Wilson, 2004).

Provincial jurisdictions, like Ontario, mandate school board improvement plans that contain an emphasis on large-scale assessments as a gauge of educational quality in both elementary and secondary schools (Volante & Ben Jaafar, 2008). For example, in their analysis of 62 Ontario school board improvement plans developed in 2003-2004, van Barneveld, Stienstra, and Stewart (2006) found that only 31% actually made reference to classroom data. Rather, it is external testing data, gathered under the direction of the provincial testing agency – the Education Quality and Accountability Office (EQAO) – that serves as the primary metric of school success. Ontario's favoritism of large-scale assessment data for driving school improvement appears, like many other jurisdictions in Canada, to be a deeply rooted practice (Klinger, Deluca, & Miller, 2008).

During the time of this study, the Ontario Ministry of Education developed an assessment, evaluation and reporting policy document for Ontario schools K-12 entitled *Growing Success: Assessment, Evaluation and Reporting in Ontario Schools*. This document, which became official policy in September 2010, aims to increase assessment literacy – an understanding of the principles and practice of sound assessment (Stiggins, 2008) – by solidifying, clarifying, and updating best assessment and evaluation practices in schools across the province. As a result of this document, common goals for schools include improved student learning, maintenance of high standards, and the formation of better mediums of communications between students, teachers, administrators and parents. Together, large-scale assessment, classroom assessment, and the *Growing Success* policy document strive to build consistency of assessment and evaluation practices across schools in the province in order to provide all students with similar academic opportunities.

The purpose of this study was to examine formative assessment practices within K-12 classrooms. More specifically, this research was guided by the following central questions:

    a) To what extent are teachers familiar with formative assessment strategies?
    b) To what extent are formative assessment strategies used within the classroom?
    c) What are the main factors that contribute to a potential gap in classroom practice?

It is important to note that these questions were part of a larger study that examined assessment literacy – an understanding of the principles and practices of sound assessment (Stiggins, 2002).

Method

Participants

Participants were selected using a mixture of purposive and convenience sampling methods across two school districts in southern Ontario, Canada. District A had an assessment consultant to support effective assessment practices within schools and s/he recruited participants by putting up a sign requesting those who were interested in an assessment study to volunteer. It seems logical that these volunteers were fairly interested in assessment issues and therefore constituted a purposeful sample (Creswell, 2008). In the other board, there was no consultant and participants were recruited through one of the professors in this study as a convenience sample. The sample consisted of 20 teachers (9 elementary, 11 secondary). Teaching experience ranged between 3 and 28 years, with a mean of 12.1. Educators were drawn from 13 schools (6 elementary, 7 secondary). Eight of the participants were male and 12 were female.

Research Site

This study was conducted in two school districts located in the Golden Horseshoe – an area around the western end of Lake Ontario, mainly the south-central region of the province. Half of the population of Ontario lives in or around this area. The student population for both districts was mixed and represented a variety of cultures and socio-economic groups.

Data Collection

The individual interviews of approximately 60 minutes involved a set of lead questions by the research team. Twenty participants were asked a range of general, open-ended questions related to assessment experience and professional development, as well as more specific questions related to their understanding and use of particular formative assessment strategies. Sample questions, some of which were piloted in a previous smaller scale study, included:
- What does formative assessment mean to you and what does it look like in your classroom?
- How do you utilize EQAO assessment results for school improvement planning?
- Please share your professional development experience in assessment and evaluation?

Each of the open-ended questions was accompanied with a set of probes designed to elicit detailed responses. For example, participants were asked to describe their use of questioning techniques, feedback without grades, self-assessment, peer assessment, and the formative use of summative assessment when answering the first question above.

Data Analysis

Analysis of the interviews followed a constant comparison approach that aligns with a recursive analysis approach (Creswell, 2008). Codes were assigned to each line

directly in the margins of the transcripts. This process was repeated for each of the transcripts. Codes from the first transcript were carried over to the second transcript, and so on. This allowed the researchers to note trends across participants. Once the initial analysis was completed, the researchers merged codes with similar meanings to create a core theme. For example, responses that detailed verbal and/or written feedback logically collapsed into one common theme: feedback without grades. Validity of the research findings was determined through cross-checking codes across the lead researchers, member check of the transcripts, clarification of the researchers' biases, and the inclusion of discrepant information (Creswell 2008). Both lead researchers analyzed the data separately and confirmed codes as a team.

Results

The results of this study identified how classroom teachers understood and used formative assessment within their classrooms. For ease of comprehension, the results are generally organized according to the theoretical framework underpinning the study. It is important to point out at the onset, that for the most part, the patterns reported could not be traced back to a particular teaching and/or training background. That is, educators who reported difficulties using particular formative assessment strategies did not come from a specific curriculum and instruction focus (e.g., mathematics, science, language arts, special education, etc), division (a division might include primary: grades 1-3, junior: grades 4-6, intermediate: grades 7-10, or senior: grades 11-12), or have a common set of professional development experiences.

Questioning

Questioning is one of the most powerful ways of "tracking student progress on a daily basis" (Elementary Teacher). The data in this study revealed a wide range of similarities in terms of teacher questioning techniques in order to improve student construction of knowledge. A secondary teacher described his old ways of questioning and how he would query students:

> Is that clear? Does everyone understand? It's as though I was asking for all the dumb kids to put up their hands. What student is going to put up [his/her] hand and confess that they don't understand when it seems like everyone else in the class gets it?

The same respondent later shared the following comment: "What questioning is really about is infusing [questioning] into daily lessons. Proper questioning techniques alleviates tension. Good questioning is really about the ability to recognize when the quiet kid doesn't get it."

Teachers in our study also argued for particular questioning techniques that are supported in the broader educational literature. For example, a secondary teacher described the model that guides questioning in her classroom:

We looked at questioning in terms of a hierarchy where the basic skills are at the bottom and higher order questioning skills are at the top of a pyramid [Bloom's taxonomy].… At the secondary level we have Q-charts [questioning techniques based on Bloom's Taxonomy] to guide our questioning.

Consider another response:

Q-charts came from the elementary level. It starts to make you think about how we check for understanding in the classroom. I can't remember thinking about my questioning techniques in my first 15 years of teaching and it wouldn't have crossed my mind that it could be used as an assessment technique. (Secondary Teacher)

Interestingly, the use of Bloom's taxonomy was never explicitly noted by elementary teachers. Collectively, the pattern of responses indicated that both elementary and secondary teachers extensively use questioning techniques in an appropriate fashion to inform their instruction and planning.

Feedback without Grades

Responses from teachers in our study suggest that a high value is placed on providing feedback without grades to students on a regular basis. Reflecting on the various formative assessment strategies, an elementary teacher noted, "I think we have made the most difference in student progress with feedback." This notion is expanded upon by another elementary teacher:

I usually give feedback without grades, because I teach grade 2 and grades don't mean much to them. Unless they have an older sibling or parent at home who's really focused on grades. For me, it's more about how can you make this better, as opposed to how can you make it an A.

A secondary teacher offers a similar response:

I teach grade 9/10 essential math [to vocational students] and I decided to break down the graphic assignment into steps to fit within the formative paradigm. If they want to resubmit an assignment seven times before the due date, then I will give them seven sources of feedback. If I give students feedback prior to when it's due, then I know that they're actually learning.

Overall, teachers in both elementary and secondary panels noted the importance of providing students with feedback that did not specifically serve an evaluative purpose.

Teacher responses also indicated a competing tension between their pursuit of a feedback-driven environment and the pragmatics of the assessment and evaluation process:

We have two forces pulling us in different directions – more assessment and less evaluation is running up against reporting more frequently for parent satisfaction and student motivation. We, as teachers, are responsible for reconciliation of various assessment tools whereby magically feedback turns into a mark for reporting. In addition, I have a hard time believing that the final assignment mark is valid: Is it their work I am marking or mine? (Secondary Teacher)

This tension between formative and summative assessment is widely acknowledged in the existing research literature (Harlen, 2005).

Another point of contention with providing too much feedback to students is raised by another secondary teacher:

I tell my students, I don't pull your mark out of the air… it's based on certain criteria and this is how your mark was added up. I put all these marks on your essay explaining what you did well, what you didn't do well, and things you can improve on, and then we give it back. [The students] just look at the mark and don't look at all the squiggles [i.e., the feedback] and then it goes into their folder or locker and [they] never pull it out again. So that doesn't really help the students; it shows [the teacher] where [students] are but unless students use it in a constructive way it's useless and a teacher's waste of time.

For the most part, teachers in this study struggled with finding creative ways to make their students fully use their formative feedback.

One secondary teacher offered the following as a way to offset challenges associated with feedback without grades:

I don't allow for rewrites in my class, instead I rely on self-assessment. I have what's called a make-up essay rather than a rewrite essay, because otherwise I am just marking my own work. So instead what I do is kind of neat. … Students must take their essay and consider the weakest part and tell me three ways that [they] can improve. … So there's lots of analysis involved and then they rewrite that segment. This way the onus is on the student to improve… and next time their essay has the potential of being better because they have learned something.

The above comments suggest teachers in our study see the value in enhancing feedback techniques and a shift in emphasis away from the final product to the process.

Self-Assessment

Despite the discomfort many educators feel towards self-assessment, the consensus among teachers from our study appears to be that involving students in the assessment process is vital to student learning. Consider the following comments from the elementary panel:

Our school is going to be focusing more on assessment, particularly assessment *as* learning [subset of formative assessment focused on meta-cognition **(Earl,**

2003)]. Of the three [purposes of assessment – assessment *of*, *for*, *as* learning], that's going to be the most important. The better the students understand what they need, the better they know what to work on. (Elementary Teacher)

A child needs to understand where he or she is having difficulties. The teacher and the student should be working together in order for it to be a learning experience for all. … We have to figure it out together. (Elementary Teacher).

It is important reiterate that assessment *as* learning is considered a subset of formative assessment that focuses on student meta-cognition. Elementary teachers in this study tended to note the importance of this assessment phase when discussing self-assessment.

Teachers in this study shared the understanding that formative assessment performed solely by the teacher is missing an integral component whereby students reflect and take ownership of their own learning:

A frequently asked question in my classroom is, "Why did I get a level 2? [on a four point scale – level 4 being the highest]. I tried really hard on that." So I decided to involve the students more in the evaluation process. After marking CASSIE [standardized reading and writing assessment], I had each of the students justify why they got the mark they did. (Elementary Teacher)

In music, self-assessment is a daily skill. It is inherent in learning to play an instrument. Self-assessment is about having enough self-criticism to say, "How can I make this better?" I guess it helps that in music, there is instant feedback [i.e., sound from an instrument]. (Secondary Teacher)

These responses suggest participants in our study reported using formative strategies that included a variety of teacher- and student-directed activities.

Teachers also acknowledged that self-assessment must be carefully implemented in order to be effective.

A teacher can't rely on self-assessment alone. Self-assessment must be preceded by an introduction from the teacher because the child might get from point A to point D but still have the wrong answer. It must be a lockstep process. (Elementary Teacher)

I think a lot of times [students] just don't know what their next steps are. Sometimes I will have my students write a list of areas in which they need to improve. I don't think it really improves their current work but it gives them a goal for next time. (Secondary Teacher)

Despite the previous responses, many teachers candidly admitted that they needed to do a much better job in promoting self-assessment within their classrooms.

Peer Assessment

Despite literature supporting peer assessment as an important formative assessment method, teachers in this study noted difficulties in the use of peer assessment and its practical application:

> The difficulty I have with formative assessment is the peer assessment portion. I just find students for whatever reason cannot be objective, or at least cannot achieve the level of objectivity that I would like. There are friendships: John is my buddy, I've known him since grade 1; or Kathy, she's pretty and I have a crush on her. And I just find I can't get an honest response when it comes down to numbers. Qualitative data is fine; it's the quantitative data that I don't trust. (Elementary Teacher)

Teachers in this study also viewed students' unfamiliarity with content as another barrier to making greater use of peer assessment within their classrooms. "A lot of the time the dilemma of the peer assessor is that s/he may not know anything about the content and might be in the interest of the student. I find peer assessment very difficult to properly implement" (Secondary Teacher). Overall, this study identified teachers' frustrations with peer assessment, with only some teachers regularly using this strategy – typically in select curriculum areas such as music and the arts.

Formative use of Summative Assessment

Both elementary and secondary teachers suggested they were able to use summative assessment, such as the provincial achievement tests, to inform their instructional practice. That is, teachers suggested they changed how they supported their students on the basis of test results. As one elementary teacher stated, "We look through those EQAO results and take them to see what areas we did poorly in and then we build goals from there." Reminiscent of the other comments, another elementary participant suggested,

> When we get the results we sit down as a division [a division might include primary: grades 1-3, junior: grades 4-6, or intermediate: grades 7-10] and have a look at where we have been and where we are going…. And that is where you can draw the conclusions and base our school growth plan on...that is pretty much a one-shot deal. You bring it in, you look at it…because it really is a snapshot that gives you more of a general direction. (Elementary Teacher)

It is important to note that the elementary teachers reported using standardized test scores but they unanimously suggested that they did so in isolation of other forms of student assessment data. This type of approach is not recommended by EQAO (2005) or the broader literature (Popham, 2005; Wilson, 2004).

Secondary teachers' use of provincial assessment data was similar to their elementary counterparts. Secondary teachers reviewed students' baseline profiles as they were reported by the test results. One participant's response was typical of the others

from this cohort: "We go to the data [and identify the specific students] who were unsuccessful." These participants used the external assessment data to create a synopsis of students' needs. In this context the literacy test "pinpoints areas of the testing where our students have problems…. For example, our students have trouble making inferences and with simple things like multiple-choice questions." Another secondary teacher commented: "We have actually taken the grade 9 math results and looked at each question specifically as to what students had difficulty with, and how we can improve." Overall, both elementary and secondary teachers recognized the utility of analyzing students' specific responses on the test itself and using this information for planning purposes. Nevertheless, responses from teachers in both elementary and secondary panels suggested they often analyzed provincial assessment results in isolation of other forms of student data (i.e. classroom assessment tasks).

Professional Development

Our study indicated that teachers often begrudge top-down, mandated professional development and do not hold much value in its execution. When specifically asked about professional development in assessment, most teachers were inclined to resist change coming from external sources. A secondary teacher explained his dislike of in-service professional development:

> It seems as though much of our PD at school is [poorly done]. Let's put all the staff in a room because we have to do something that we can go back to the SOs [Superintendants Office] and say that we did what he expected of us. So we are told to read this article and think of some new ideas, but not really because when do they listen to us teachers anyway? Nothing significant happens afterwards. You know, I think it all goes in the garbage when we're done. I think it would be more productive if we just twiddled our thumbs.

Although many of the respondents did not have the same tone as the previous teacher, many echoed their frustration with the professional development provided by their board.
Teacher responses tended to underscore the importance of self-directed professional development approaches.

> About 5 years into my career, I began an Action Research Project developing portfolio assessment. Since then, portfolio assessment has been taken on by the board and so the portfolio continues to be an important part of our assessment in order to support student learning. For my personal professional development, I guess you could say I do a little bit of reading but mostly my PD comes from PLCs [Professional Learning Communities] focusing on assessment; we have done moderated marking of writing assessments. (Elementary Teacher)

Overall, participants in our study indicated a strong preference for self-directed approaches to professional development in order for more sustained changes to be realized in their classroom practice.

Our study also indicated that the primary source of professional development for these teachers came from a faculty of education – either from a teacher education program or at the graduate level. One secondary teacher said, "I took some assessment and evaluation courses through my M.Ed… There is a lot more in-service that is being offered at the board level but I still rely on what I learned in my masters." Another elementary teacher agreed with the previous statement, "I think my master's courses offered the most information in terms of assessment. I still pull out my textbook from time to time." A few teachers in our study commented on the value of coupling university education with in-service professional development, specifically moderated marking sessions. Consider the following response:

> I've taken my reading specialist and a University-type course in A&E [assessment and evaluation] but that was awhile back… Recently our school has initiated a lot of moderated marking with other schools… we all sit down together and define the criteria and determine what's proficient, what's exemplary… this way we're all doing the same thing. (Elementary Teacher)

Building consistency within assessment practices, across not only subject areas but also grades and schools, was a primary purpose of many professional development sessions. The elementary and secondary teachers in our study agreed that professional development should continue to aid in teachers' understanding and use of daily assessment practices to improve teacher practice and student learning.

Discussion

More than ever, teachers are required to be accountable (e.g., standardized testing, curriculum expectations, evidence-based practice), while they simultaneously negotiate an unprecedented level of student involvement in the assessment process. Our findings indicated that teachers are becoming more familiar with a diverse range of formative assessment strategies and are reporting their use on a more consistent basis within their classrooms. Given the size of our sample, it would be instructive to examine the extent of these findings in other education jurisdictions.

Nevertheless, the present study also indicated that there was an imbalance in the use of particular formative assessment strategies associated with improvements in student learning (Black & Wiliam, 1998; Popham, 2005; Stiggins, 2008). When considered along with the fact that many participants in this study were nominated by board representatives for their interest in classroom assessment, the present results likely overestimate the degree to which particular formative strategies such as peer and self-assessment are actually being used. Thus, targeted professional development and greater attention at the pre-service and in-service level seems warranted by the present results.

It can also be understood from this study that teachers are beginning to value learning over a singular focus on grades; despite the constant student and parent pressures they feel around providing more evaluative reporting. Although evaluation of student work is still important for communication and reporting purposes, on a day-to-day basis, teachers in our study tended to emphasize feedback without grades rather than the final evaluation. The only drawback to this approach is the question of the validity of student

work. As one secondary teacher commented "Is it their [students'] work I am marking or mine [teachers]?" This implies that feedback without grades must be used in a meaningful way, focusing on the learning that is taking place rather than the final grade. Negotiating the pragmatic challenges of providing feedback versus evaluation continues to be an area worthy of future professional development, particularly at the secondary level.

One of the other pivotal findings from this study is that assessment should be a collaborative process, including the teacher, student, and peers. At times, it can be difficult for teachers to relinquish authority (which is often seen in the power to give out marks), but teachers from this study indicate that the ability for students to "self-criticize" is key to understanding why a particular grade was received as well as for self-improvement purposes. As one of the participants noted, "It is a lockstep process" whereby teachers provide guidance while students attempt to self-assess and reflect on what they have learned. Students are also able to see in others' work what they have omitted from their own. Interestingly, the minority of teachers that were making use of peer assessment recommended that student judgments be confined to qualitative feedback to avoid outside factors interfering with the overall assessment (i.e. friendships, lack of content knowledge, interest, etc.).

Despite most teachers' negative connotations with summative large-scale assessment, almost all teachers in this study used EQAO results in some capacity. The most common uses were forming baselines of student performance and highlighting trends across grades and years. This, in turn, provided schools and classrooms with focal "big ideas" and consequently starting points for instruction, yielding a classroom of evidence-based practice. Nevertheless, the findings also suggested that teachers continue to grapple with synthesizing the results of provincial assessments with other salient forms of student data (i.e., classroom assessments, other standardized tests, English-as-a-Second-Language status, etc). Despite probing, the majority of participants were not successful in the integration and implementation of such practices. Since research overwhelmingly supports the relationship between instructional improvement and the prudent use of different forms of student data (Earl & Torrance, 2000; Heritage & Chen, 2005; Sutherland, 2004; Timperley, 2005; Wohlstetter, Datnow, & Park, 2008), it appears that greater attention must be directed to equipping teachers with this type of instructional capacity.

One of the more interesting findings from our study was that many educators stated their primary source of professional development was from a faculty of education – either from a teacher education program or at the graduate level. Yet it is important to recognize that few faculties of education teach courses on assessment and evaluation. For instance, Klinger (2009) noted that out of 18 teacher education programs in Ontario, only two universities offer a separate course in classroom assessment, while the other programs embed assessment into 'teachable' subject areas such as mathematics, science or English. The limitation of this design, according to Klinger (2009), is that not all faculty members have expertise in assessment and evaluation and, therefore, the assessment content is not infused properly. Thus, teacher education reforms are pivotal for improving teacher competence in all facets of formative assessment. Ultimately, when jurisdictions create the conditions for educators to learn new skills, knowledge, attitudes,

and beliefs, an increase in consistency across teachers' best practices can be more fully realized (Stoll, 2009).

Although previous research has noted differences in the use of summative assessment methods across elementary and secondary schools (Volante, 2010), the present findings suggested there were few, if any, differences across panels with respect to formative assessment. The only exception to the previous trend was the greater use of Bloom's taxonomy for guiding questioning techniques at the secondary level and a tendency for secondary teachers to report more tensions with providing feedback without grades. The more substantive trend, however, was that elementary and secondary teachers noted difficulties in the effective use of particular formative strategies such as peer and self-assessment. Essentially, formative assessment must be treated as a multi-faceted concept with greater attention given to the pragmatic challenges associated with particular techniques. In doing so, teacher educators can improve the quality of their professional development programs and foster a more balanced assessment framework in K-12 schools.

Conclusion

As Ontario continues to move forward with the implementation of its new assessment and evaluation policy framework *Growing Success* (Ontario Ministry of Education, 2010), it is essential that the policy implementation strategy carefully consider teachers' perceptions and self-efficacy in particular facets of classroom assessment. As previously stated, teachers' perceptions strongly influence how they teach and what their students ultimately learn (Brown, 2004). Findings from our study are consistent with previous research pertaining to formative assessment and suggest that teachers, for the most part, are using particular formative assessment practices such as questioning techniques, feedback without grades, and the formative use of summative assessment techniques with a moderate level of success. Alternatively, there are some practical barriers associated with peer and self-assessment that affect teachers' self-efficacy and their willingness to fully execute such practices. Generating coherence between ministries of education, faculties of education, school boards, and individual schools is essential for sustainable reform. It is our hope that this study will act as a catalyst for greater attention to the conditions and factors that foster a balanced classroom assessment approach. Ultimately, greater synergy between formative assessment research and classroom practice remains a key priority for enhancing student learning.

References

Assessment Reform Group. (2002). *Assessment for learning: 10 principles*. Retrieved from http://www.assessment-reform-group.org/publications.html

Black, P., Harrison, C., Lee, C., Marshall, B., & Wiliam, D. (2004). Working inside the black box: Assessment for learning in the classroom. *Phi Delta Kappan*, *86*(1), 9-21.

Black, P., & Wiliam, D. (1998). Inside the black box: Raising standards through classroom assessment. *Phi Delta Kappan*, *80*(2), 139-148.

Black, P., & Wiliam, D. (2006). Assessment for learning in the classroom. In J. Gardner (Ed.), *Assessment and Learning* (pp. 9-14). Thousand Oaks, CA: Sage.

Brown, G. T. L. (2004). Teachers' conceptions of assessment: Implications for policy and professional development. *Assessment in Education*, *11*(3), 301-318. doi:10.1080/0969594042000304609

Creswell, J. W. (2008). *Educational research: Planning, conducting, and evaluating quantitative and qualitative research* (3rd ed.). Upper Saddle River, NJ: Merrill Prentice Hall.

Crooks, T. J. (1988). The impact of classroom evaluation practices on students. *Review of Educational Research*, *58*, 438-481. doi:10.3102/00346543058004438

Decker, D. M., & Bolt, S. E. (2008). Challenges and opportunities for promoting student achievement through large-scale assessment results. *Assessment for Effective Intervention*, *34*(1), 43-51. doi:10.1177/1534508408314173

Dekker, T., & Feijs, E. (2005). Scaling up strategies for change: Change in formative assessment practices. *Assessment in Education*, *12*(3), 237-254. doi:10.1080/09695940500337215

Earl, L. (2003). *Assessment as learning*. Thousand Oaks, CA: Corwin.

Earl, L., & Katz, S. (2006). *Rethinking Classroom Assessment with a Purpose in Mind*. Retrieved from http://www.edu.gov.mb.ca/k12/assess/wncp/rethinking_assess_mb.pdf

Earl, L., & Torrance, N. (2000). Embedding accountability and improvement into large-scale assessment: What difference does it make? *Peabody Journal of Education, 75*(4), 114-141. doi:10.1080/09695940802164226

EQAO (2005). *Using data to promote student success (Grades 3 & 6 assessments of reading, writing and mathematics)*. Toronto, Ontario: Queen's Printer of Ontario.

Harlen, W. (2005). Teachers' summative practices and assessment for learning –
    Tensions and synergies. *The Curriculum Journal*, *16*(2), 207-223.
    doi:10.1080/09585170500136093

Harlen, W. (2007). *Assessment of Learning*. Thousand Oaks, CA: Sage.

Heritage, M., & Chen, E. (2005). Why data skills matter in school improvement. *Phi
    Delta Kappan, 86*(9), 707-710.

Khan, E. A. (2000). A case study of assessment in a grade 10 English course. *The Journal
    of Educational Research*, *93*, 276-286.

Klinger, D., Deluca, C., Miller, T. (2008). The evolving culture of large-scale
    assessments in Canadian education. *Canadian Journal of Educational
    Administration and Policy*, (76). Retrieved from
    http://umanitoba.ca/publications/cjeap/articles/klinger.html

Klinger, D. (2009, April). Developing a curriculum for assessment education. Paper
    presented at the American Educational Research Association Conference, San
    Diego, CA.

Kluger, A. N., & DeNisi, A. (1996). The effects of feedback intervention on
    performance: A historical review, a meta-analysis, and a preliminary feedback
    intervention theory. *Psychological Bulletin*, *119*(2), 254-284. doi:10.1037/0033-
    2909.119.2.254

Natriello, G. (1987). The impact of evaluation processes on students. *Educational
    Psychologist*, *22*(2), 155-175. doi:10.1207/s15326985ep2202_4

Ontario Ministry of Education. (2010). *Growing Success: Assessment, Evaluation, and
    Reporting in Ontario's Schools*, *First Edition Covering Grades 1-12*. Toronto,
    ON: Queen's Printer for Ontario. Retrieved from
    http://www.edu.gov.on.ca/eng/policyfunding/growSuccess.pdf.

Popham, W. (2005). *Assessment for Educational Leaders*. Upper Saddle River, NJ: Allyn
    & Bacon.

Popham, W. J. (2009). *Instruction that measures up: Successful teaching in an age of
    accountability.* Alexandria, VA: ASCD.

Stiggins, R. (2002). Assessment crisis: The absence of assessment for learning. *Phi Delta
    Kappan*, *83*(10), 758-765.

Stiggins, R. J. (2008). *Student-Involved Assessment for Learning* (5th ed.). Upper Saddle
    River, NJ: Prentice Hall.

Stoll, L. (2009). Capacity building for school improvement or creating the capacity for learning? A changing landscape. *Journal of Educational Change*, *10*, 115-127. doi:10.1007/s10833-009-9104-3

Sutherland, S. (2004). Creating a culture of data use for continuous improvement: A case study of an Edison Project School. *American Journal of Evaluation, 25*(2), 277-293. doi:10.1177/109821400402500302

Timperley, H. S. (2005). Instructional leadership challenges: The case of using student achievement information for instructional improvement. *Leadership and Policy in Schools, 4,* 3-22.

van Barneveld, C., Stienstra, W., & Stewart, S. (2006). School improvement plans in relation to the AIP model of educational accountability: A content analysis. *Canadian Journal of Education*, *29*(3), 839-854.

Volante, L. (2010). Assessment of, for, and as learning within schools: Implications for transforming classroom practice. *Action in Teacher Education*, *31*(4), 66-75.

Volante, L. (2008). Equity in multicultural student assessment. *Journal of Educational Thought*, *42*(1), 11-26.

Volante, L., & Ben Jaafar, S. (2008). Profiles of educational assessment systems worldwide: Educational assessment in Canada. *Assessment in Education*, *15*(2), 201-210. doi: 10.1080/09695940802164226

Wiliam, D. (2007). Content then process: Teacher learning communities in the service of formative assessment. In D. B. Reeves (Ed.), *Ahead of the curve: The power of assessment to transform teaching and learning* (pp. 183-204). Bloomington, IN: Solution Tree.

Wilson, M. (Ed.). (2004). *Towards coherence between classroom assessment and accountability: 103rd yearbook of the National Society for the Study of Education, Part II*. Chicago: University of Chicago Press.

Wilson, N. S. (2008). Teachers expanding pedagogical content knowledge: Learning about formative assessment together. *Journal of In-Service Education*, *34*(3), 283-298. doi:10.1080/13674580802003540

Wohlstetter, P., Datnow, A., & Park, V. (2008). Creating a system for data-driven decision-making: Applying the principal-agent framework. *School Effectiveness & School Improvement*, *19*(3), 239-259. doi:10.1080/09243450802246376