Changing practices: influences on classroom assessment

Robin D. Tierney*
University of Ottawa, Canada

The pedagogical potential of classroom assessment to support student learning has increasingly been evidenced in research over the past decade. Constructive classroom assessment has been championed by assessment specialists, and endorsed by professional organizations. In practice, however, the process of changing classroom assessment from its traditionally summative orientation is not straightforward. This methodical review looks at how six sources, which are educational research, evaluative inquiry, large-scale assessment, educational policy, professional development, and teachers’ beliefs, influence and mediate assessment practices. A group of purposively selected research articles are analysed as evidence of the dynamics in this complex process. Cross-currents relating to research perspective, collaboration, and time are discussed. For the movement seen in this study to continue, the tension between teacher autonomy and school community, and the relationship between collective commitment and assessment literacy should be considered.

The use of classroom assessment to promote student learning is strongly supported by current educational research. Following the seminal review by Black and Wiliam (1998), a host of empirical work has confirmed the pedagogical potential of classroom assessment (e.g., Black & Harrison, 2001; Barootchi & Keshavarez, 2002; Ormond et al., 2002; Coffey, 2003; Lee & Gavine, 2003; Waddell, 2004). Despite this research evidence, the sustained championing by assessment specialists (e.g., Stiggins, 1994, 2001; Stiggins & Chappuis, 2005), and the increasing endorsement by professional organizations (Joint Committee on Standards for Educational Evaluation, 2003; Miller, 2005), assessment is still not widely used to promote learning in elementary and secondary classrooms. On the contrary, the summative function is emphasized, and teachers continue to use classroom assessment primarily for grading and reporting (Kehr, 1999; McNair et al., 2003; Uchiyama, 2004). Although the rhetoric for assessment reform is strong, the way in which student learning is assessed in classrooms on a regular basis seems resistant to change.

*Faculty of Education, University of Ottawa, Lamoureux Hall, 145 Jean-Jacques Lussier, Ottawa, Ontario, Canada K1N 6N5. Email: r tier017@uottawa.ca
Educational change is depicted as both slow in pace, and recursive in nature. Cuban (1990) compares the process of reform to ocean waves, and Hayward et al., (2004) elaborate on this metaphor in an article on the connection between educational policy and practice. They compare educational reform to a hurricane that whips up waves across the ocean surface without disturbing the essential teaching and assessment practices that lie below. This paper is concerned with how assessment practices are influenced by six sources of information in education, which are research, evaluative inquiry, large-scale assessment, policy, professional development, and teachers’ beliefs. Metaphorically, the focus here is less on the frenetic waves of reform, and more on the estuaries, where currents meet and ever-changing patterns are created in the residue.

The process of educational change has already been the subject of considerable research (e.g., Brighton’s 2003 review), but the literature tends to consider change on a broad scale as a systemic or organizational challenge (Horn, 2002; Fullan, 2003). This paper deals with change at the classroom level. The disparity between current knowledge about classroom assessment and common practice in elementary and secondary classrooms calls the influence of educational information into question. How are sources of information in the field of education influencing change toward constructive practices in classroom assessment? To respond to this question, a methodical review and analysis of recent journal articles was undertaken, with a focus on the change process and dynamics in studies relating to classroom assessment. The balance of this article explains the conceptual framework and methodology, describes the findings, and discuses currents and tensions that were seen through the analysis.

**Conceptual framework**

**Research perspective**

The mode of inquiry for this work is best described as critical pragmatism. From this perspective, knowledge is considered socially constructed, contextual, and impermanent (Biesta & Burbules, 2003; De Waal, 2005). Critical pragmatism is well-suited to educational research because it focuses on practice (action) and interaction.

**Sources of influences**

A model with seven main elements was created as an organizational tool for this study (see Figure 1). Informed by literature on the use of information in educational reform (Cousins & Leithwood, 1993; Fullan, 1993; Horn, 2002), the model includes six sources of influence. Three of these, (1) educational research, (2) evaluative inquiry, and (3) large-scale assessment, are considered knowledge-generating sources. Evaluative inquiry and educational research are superficially similar in that they both involve a process of inquiry in which evidence is gathered, and they are both epistemologically and methodologically diverse. Their intent, however, is different.
Educational research is designed to contribute to a knowledge base, and to further understanding for both practical and theoretical purposes about a particular area of scholarly interest. Evaluative inquiry is also concerned with understanding, but in terms of the merit, worth, and significance of an object in relation to specified criteria (Cousins et al., 2004). As a result of this fundamental difference, evaluative inquiry and educational research are considered here as two distinct sources of influence in the model.

The third knowledge-generating source, large-scale assessment, is also a form of educational inquiry as it attempts to determine and describe the achievement of large numbers of students at a given point in time. The combination of technological advancements and political impetus in the later part of the twentieth century gave educational organizations worldwide the ability to collect, store, and analyse vast amounts of data. In the 1990s many national and provincial educational organizations followed the lead of the international assessment organizations, and implemented large-scale assessments of student achievement. The use of data for accountability purposes is well established, but a more potent role exists for low-stakes data as a source for educational improvement (Earl, 2001; Simon & Forgette-Giroux, 2002; Earl & Fullan, 2003; Ingram et al., 2004; Sutherland, 2004). The
possibility of large-scale assessment as a positive source of influence on classroom practices stems from this view.

In constructing the framework model, it was thought that information would not flow directly from the three knowledge-generating sources to the classroom, but would be mediated by (4) educational policy, (5) professional development, and (6) teachers’ beliefs and conceptions. The basis for this idea lies in the research relating to teachers’ perceptions about, or use of various sources of information (Cousins & Leithwood, 1993; Landrum et al., 2002; Ontario College of Teachers, hereafter, OCT, 2003). Although some teachers are actively involved in educational inquiry, recent survey results indicate that many place little value on university courses and research journals, where academic information is disseminated (Landrum, et al., 2002; OCT, 2003). Colleagues are most valued as a source of information for teachers, but professional development is also considered important (OCT, 2003), and a preference for local in-service is suggested (Cousins & Leithwood, 1993). Whereas educational policy and professional development are external sources of influence, teachers’ beliefs and conceptions are considered an internal mediator, and their interplay with classroom assessment practices is well established (Delandshere & Jones, 1999; Gilsdorf, 2000; McMillan & Nash, 2000; Schmidt, 2001; McMillan et al., 2002). The arrows in the model indicate the potential currents of change in a broad educational context, with (7) assessment practices in elementary and secondary classrooms as the assumed point of confluence. Classroom practice may inform all types of educational inquiry; therefore, a final arrow indicates a possible return from the classroom to the three knowledge-generating sources.

**Methodology**

**Overview**

This study draws methodological guidance from several sources on systematic reviews (Davies, 2000; Evans & Benefield, 2001; Hammersley, 2001), and from the work of Sandelowski et al. (1997) on issues in qualitative metasynthesis. The term methodical review is used to differentiate the method from both meta-analysis and meta-synthesis. While being thorough and orderly, the research process used here does not have the quantitative goals (e.g., effect sizes) that are often associated with meta-analysis (Hammersley, 2001), nor does it attempt to synthesize research results (Sandelowski et al., 1997). The 24 articles studied in this paper are considered as evidence of a complex process whose dynamics can be better understood by highlighting common currents and discussing divergences.

**Data sources: selection criteria**

The importance of making the selection criteria for reviews explicit is stressed by several authors (Sandelowski et al., 1997; Davies, 2000; Evans & Benefield, 2001), but also criticized as a futile attempt to eliminate subjectivity (Hammersley, 2001).
Here it is viewed pragmatically as a means of making methodological decisions, regardless of their degree of subjectivity or objectivity, as transparent as possible. To begin the selection process, 280 articles relating to classroom assessment were identified using 29 Boolean search combinations in the ERIC and Scopus databases. This comprehensive manner of searching was necessary given that terminology in this area is inconsistent, and assessment practices that promote student learning are sometimes referred to as formative, embedded, interactive, constructive, dynamic, or responsive assessment, as well as assessment for learning, or assessment as learning.

Four types of selection criteria were determined for the initial search: document type, publication date, context, and focus. All of the retained articles were published in refereed scholarly journals between January 2000 and March 2005, and they all offer empirical evidence, either in the form of a first-hand report based on observation (e.g., Boardman & Woodruff, 2004), or as an attempt to understand existing research from alternative angles (e.g., Tunstall, 2001; Hargreaves et al., 2002). The 24 studies selected for review are all concerned with changing classroom assessment practices. The context was limited to elementary and secondary classroom education, which excludes post-secondary adult education, pre-service teacher education, early childhood, or preschool education, home schooling, and distance education. Studies relating to individual psychological assessment, or computer-based testing were also excluded as this study focuses on assessment as an interaction between teachers, students, and other stakeholders in the social process of classroom learning.

Data analysis

A systematic process that involved several readings of each article was employed. Three working tables were created based on the content of the selected articles. The first working table contains information about the purpose of each study, the key concepts, and the underlying epistemology. The second working table focuses on methodology, identifying the method for each study, as well as evidence of collaboration, participants and contexts, data sources, and the research process. In the third working table key findings are summarized, the role of teachers is identified, along with barriers or facilitators in the change process and any impact on classroom assessment practices. To understand the flow of influence as it was seen within each study in the sample, a diagram was created for each article in relation to this study’s conceptual framework.

Findings

Overview

The six knowledge-generating and mediating sources included in the conceptual model were all seen in the sample. The strength of showing was determined by the frequency of association, meaning simply that a count was done of the number of articles relating to each source. The majority can be linked to three or more sources,
and all are associated with educational research to some degree. Fewer articles are associated with the other two knowledge-generating sources, with nine being related to large-scale assessment, and only four with evaluative inquiry. Stronger links are made with the mediating sources of influence: 14 are associated with professional development, 17 with educational policy, and 19 with teacher beliefs. While the association frequency indicates the degree to which each source of influence is present in the sample, it tells little about the interaction between sources. The flow diagrams that were created for analysis do illustrate this activity, but considered as a whole, their complexity defies two-dimensional modelling. Consequently, the articles are described first in six groups based on the conceptual framework, and then discussed thematically. Key information on each article is distilled in Table 1 for quick reference.

Knowledge-generating sources

Educational research. All of the studies in the sample relate to university-based research in that they are all published in research journals, and they all include some form of literature review. Given the medium, this universal use of educational research is not surprising. However, a notable range is seen in the extent to which research is used, and in the role it plays in these studies. Almost all (20 of 24) of the authors refer to at least one methodological influence, and several provide numerous citations for their research process (e.g., Tunstall, 2001; Kyriakides, 2002; Mabry et al., 2003). Most often, existing literature is used for justification, either as a brief rationale for the work (e.g., Yung, 2001; Hall & Harding, 2002), as background information for a case (Borko et al., 2000; Lock & Munby, 2000; Treagust et al., 2001), or to explain a particular strategy (e.g., Swanson & Stevenson, 2002). Research is often reviewed as part of a conceptual or theoretical framework, and in some cases, the review is extensive (Hand & Prain, 2002; Kyriakides, 2002; Morgan & Watson, 2002; Taylor et al., 2005). Other researchers are inspired by a single source, and they adapt existing models to understand classroom assessment in reform contexts (Hall, 1977 as cited in Cheung, 2002; House, 1981 as cited in Hargreaves et al., 2002).

Another commonality is seen in the use of existing work to reach new conclusions. In two articles, data previously collected by the authors is viewed through different theoretical lenses. Hargreaves et al. (2002) re-examine teachers’ ideas about assessment in Ontario from four perspectives: the technological, cultural, political, and postmodern. Tunstall (2001) re-examines assessment discourse, originally considered using grounded theory (Tunstall & Gipps, 1996), from a critical poststructuralist perspective. Torrance and Pryor (2001) also build on their own research by taking theory developed in an earlier project on formative assessment (Torrance & Pryor, 1998) back to the classroom. McDonald’s (2002) study is somewhat unique in this subset in that it builds longitudinally on an earlier study, using replication to measure change in educators’ attitudes. A key observation about this group of articles is that none link educational research directly to classroom practices, and in all 24 articles,
<table>
<thead>
<tr>
<th>Author (Date)</th>
<th>Purpose</th>
<th>Method &amp; Data Sources</th>
<th>Participants &amp; Context</th>
</tr>
</thead>
<tbody>
<tr>
<td>Boardman &amp; Woodruff</td>
<td>To determine how state accountability tests influenced use of an instructional approach introduced through professional development sessions relating to CSR, an instructional (reading) method. Part of a larger study on different professional development approaches.</td>
<td>Qualitative: Grounded theory using observation field notes, teacher implementation logs, interview transcripts and field notes.</td>
<td>20 elementary teachers in 3 schools in urban district in accountability context in Texas, USA.</td>
</tr>
<tr>
<td>Borko, Davinroy, Bliem &amp; Cumbo (2000)</td>
<td>To examine the process of teacher change, and factors that facilitate and hinder it, in the context of a reform-based collaborative staff development project.</td>
<td>Qualitative: Collaborative case study using workshop field notes, student artefacts, interview transcripts, observation field notes.</td>
<td>2 experienced elementary teachers selected purposively from earlier professional development project in Colorado, USA.</td>
</tr>
<tr>
<td>Brighton (2003)</td>
<td>To identify and describe middle school teachers' beliefs about teaching and learning in diverse classrooms, and to determine how those beliefs affect teachers' will and ability to differentiate their teaching and assessment practices. Part of a larger study.</td>
<td>Qualitative: Systematic grounded theory using interview transcripts, classroom observation notes, documentary evidence, and student artefacts.</td>
<td>Target teachers (no #) and stratified sample of students from target classrooms (no #) in 4 middle schools in eastern USA.</td>
</tr>
<tr>
<td>Briscoe &amp; Wells (2002)</td>
<td>To examine the action research process, how a teacher overcame problems in that process, and how it affected the changes she made in her science teaching and assessment practices.</td>
<td>Qualitative: Collaborative case study using field notes, action research journal, documentary evidence, student artefacts, and action research report.</td>
<td>1 experienced elementary teacher/graduate student in Florida, USA.</td>
</tr>
<tr>
<td>Author (Date)</td>
<td>Purpose</td>
<td>Method &amp; Data Sources</td>
<td>Participants &amp; Context</td>
</tr>
<tr>
<td>--------------------</td>
<td>-------------------------------------------------------------------------</td>
<td>---------------------------------------------------------------------------------------</td>
<td>----------------------------------------------------------------------------------------</td>
</tr>
<tr>
<td>Cheung (2002)</td>
<td>To understand and provide information for educational authorities about teachers’ concerns in the expansion of the TAS, a system assessment scheme using classroom assessment, with the Stages of Concern Model developed and tested in an earlier study by the author.</td>
<td>Mixed: Survey research using 1 questionnaire with 22 Likert-scales and 1 with extended response open-ended question.</td>
<td>290 secondary school teachers in Hong Kong, of which 55 already TAS teachers.</td>
</tr>
<tr>
<td>Dori (2003)</td>
<td>To determine as part of an evaluative inquiry, the science achievement of secondary students in the 5th year of the Matriculation 2000 project, a government-sponsored effort to change classroom assessment.</td>
<td>Mixed: Experimental (control groups) used within an interpretive evaluation, drawing on achievement test data and observation visits.</td>
<td>243 grade 12 students in 4 exemplary high schools in Israel.</td>
</tr>
<tr>
<td>Hall &amp; Harding (2002)</td>
<td>To understand the extent to which communities of assessment practice exist, focusing on the mechanism and structures that supported the sharing of levelled descriptions of assessment criteria (National Curriculum scales) amongst staff and with students and parents.</td>
<td>Qualitative: Multiple cases using interview transcripts, student artefacts, and documentary evidence.</td>
<td>All grade 2 teachers and assessment coordinators, and LEA assessment advisors (no #) for 6 schools in Northern England, UK.</td>
</tr>
<tr>
<td>Hand &amp; Prain (2002)</td>
<td>To investigate key issues in the change process, and report on the changing beliefs and practices of two teachers as they engaged in a professional development programme over 4 years. Part of a larger study.</td>
<td>Qualitative: Case study approach using interview transcripts, classroom observation notes, student artefacts, field notes from group professional development sessions.</td>
<td>2 experienced secondary science teachers selected purposively from larger study in Australia.</td>
</tr>
</tbody>
</table>
Table 1. (Continued)

<table>
<thead>
<tr>
<th>Author (Date)</th>
<th>Purpose</th>
<th>Method &amp; Data Sources</th>
<th>Participants &amp; Context</th>
</tr>
</thead>
<tbody>
<tr>
<td>Hargreaves, Earl &amp; Schmidt (2002)</td>
<td>To illustrate and exemplify various perspectives in classroom assessment reform using data from an earlier study by the authors on the experiences of teachers during reform.</td>
<td>Qualitative: Secondary interpretation using interview transcripts, observation notes, and group meeting notes.</td>
<td>29 grade 7 and 8 teachers in Ontario, Canada.</td>
</tr>
<tr>
<td>Hayward, Priestley &amp; Young (2004)</td>
<td>To describe the change process and barriers for sustainability of a government-approved assessment programme in the pilot-project phase.</td>
<td>Qualitative: Interpretive analysis of evaluative reports completed by schools participating in programme.</td>
<td>Teachers (no #) in 33 elementary and secondary schools in Scotland.</td>
</tr>
<tr>
<td>Kyriakides (2002)</td>
<td>To determine whether baseline assessment can be used both for formative purposes and for value-added analysis in a study that assesses primary mathematics knowledge over 2 years.</td>
<td>Mixed: Method not explicit. Focus on test construction. Uses observation checklist, performance test results, student interviews, and achievement test data.</td>
<td>All grade 1 students (no #) in 58 schools in Cypress.</td>
</tr>
<tr>
<td>Lock &amp; Munby (2000)</td>
<td>To study a teacher’s beliefs and practices during integration of new methods into existing assessment practices. Part of a larger university-based research project aiming to collaborate with teachers regarding best classroom assessment practices.</td>
<td>Qualitative: Case study using classroom observation notes, interview transcripts, meeting notes, student artefacts, documentary evidence, and fieldwork journal.</td>
<td>1 experienced teacher in split 7/8 class in small, middle-income urban school in Ontario, Canada.</td>
</tr>
<tr>
<td>Lovett &amp; Gilmore (2003)</td>
<td>To document and describe the effectiveness of a professional development model, the Quality Learning Circle, in which teachers direct their own learning, and are active partners in the initiative. Follows earlier study by author.</td>
<td>Qualitative: Method not explicit. Uses interview transcripts, QLC meeting observation notes, and documentary evidence.</td>
<td>8 teachers invited from earlier study to be members of a Quality Learning Circle in urban New Zealand.</td>
</tr>
<tr>
<td>Author (Date)</td>
<td>Purpose</td>
<td>Method &amp; Data Sources</td>
<td>Participants &amp; Context</td>
</tr>
<tr>
<td>--------------</td>
<td>---------</td>
<td>----------------------</td>
<td>------------------------</td>
</tr>
<tr>
<td>Mabry, Poole, Redmond, &amp; Schultz (2003)</td>
<td>To determine the impact of standards-based reform and associated state testing, especially in relation to classroom practices that support student learning.</td>
<td>Qualitative: Interpretive study using interview transcripts/notes and documentary evidence related to state testing.</td>
<td>31 elementary and secondary teachers purposively selected by students and researchers in Washington, USA.</td>
</tr>
<tr>
<td>Mason, Mason, Mendez, Nelsen &amp; Orwig (2005)</td>
<td>To describe district-wide reform efforts in developing curriculum standards and student assessment with strong teacher involvement. Focus on implementation strategies, teachers’ perceptions, and effects of reform efforts on student achievement in reading and mathematics.</td>
<td>Mixed: Case study approach combined with statistical modelling. Uses interview transcripts, documentary evidence (e.g. report cards, development guidelines, etc.), reports by district leaders, and student achievement data.</td>
<td>4 district leaders who are also co-authors. Approx 18,000 students in 22 elementary and secondary schools in large district in California, USA.</td>
</tr>
<tr>
<td>McDonald (2002)</td>
<td>To investigate and compare teachers’ and administrators’ perceptions about use of diploma examination data in a replication and extension of Loerke’s (1993) study.</td>
<td>Mixed: Survey research followed by interviews. Uses questionnaire with forced-choice items and background questions and interview transcripts.</td>
<td>363 teachers and 40 school administrators in high schools and 10 interviewees in Alberta, Canada.</td>
</tr>
<tr>
<td>Morgan &amp; Watson (2002)</td>
<td>To discuss the interpretive nature of assessment and exemplify issues of fairness and equity in mathematics assessment, this paper draws on two empirical studies previously conducted independently by the authors. The 1st looks at the ‘robustness’ of judgements made by the teacher, and the 2nd concentrates on evaluations by teachers using specified criteria.</td>
<td>Qualitative: 2 independent case studies. 1st uses classroom observations, interview transcripts and field notes; 2nd uses think-aloud interviews as teachers evaluate samples of student writing.</td>
<td>1 experienced grade 7 teacher in 1st study, and 11 experienced teachers from 5 secondary schools in 2nd study. Both within UK reform context.</td>
</tr>
</tbody>
</table>
Influences on classroom assessment

Table 1. (Continued)

<table>
<thead>
<tr>
<th>Author (Date)</th>
<th>Purpose</th>
<th>Method &amp; Data Sources</th>
<th>Participants &amp; Context</th>
</tr>
</thead>
<tbody>
<tr>
<td>Swanson &amp; Stevenson (2002)</td>
<td>To investigate the relationship between policy and practice in the educational reform process by measuring first the impact of the national standards movement on policy activism across the fifty states, and second the degree of alignment between state-level standards activism and teachers’ use of standards-based practices in the classroom.</td>
<td>Quantitative: Rasch measurement analysis and hierarchical linear analysis (HLM) using NAEP mathematics achievement and teacher questionnaire data, and 1996/7 Council of Chief of State School Officers studies.</td>
<td>Longitudinal database created using 1992 and 1996 cross-national data sets with classroom as the unit of analysis. Final data set included 30 states with over 1000 schools in USA.</td>
</tr>
<tr>
<td>Taylor, Pearson, Peterson &amp; Rodriguez (2005)</td>
<td>To determine the effectiveness of a large-scale, structured professional development programme, the School Change Framework (CIERA), with a focus on classroom-level factors, as a means of school reform and improving student achievement in reading and writing.</td>
<td>Mixed: Naturalistic experiment using interview transcripts, monthly summary logs, end-of-year team reports, classroom observation notes and checklist, and 3 standardized reading tests.</td>
<td>93 teachers, 733 students in grade 2 to 5, and principals in 15 elementary schools in 5 states in USA.</td>
</tr>
<tr>
<td>Torrance &amp; Pryor (2001)</td>
<td>To consider the validity and utility of a theoretical model of formative assessment as it is integrated into classroom practice, and to understand whether collaborative action research was helpful as a means of changing classroom assessment practices.</td>
<td>Qualitative: Action research approach using reports from teacher researchers, team meeting tapes, interview transcripts, classroom observation tapes, and teacher artefacts from professional development workshops.</td>
<td>5–7 elementary teacher-researchers and 2 university researchers in UK.</td>
</tr>
<tr>
<td>Author (Date)</td>
<td>Purpose</td>
<td>Method &amp; Data Sources</td>
<td>Participants &amp; Context</td>
</tr>
<tr>
<td>--------------</td>
<td>---------</td>
<td>-----------------------</td>
<td>------------------------</td>
</tr>
<tr>
<td>Tregust, Jacobowitz, Gallagher &amp; Parker (2001)</td>
<td>To document and describe the use of embedded assessment, and its effects on students and the learning environment in classroom. Part of a larger cooperative project linking research and professional development.</td>
<td>Qualitative: Interpretive case study using lesson tapes and interview transcripts.</td>
<td>1 experienced teacher in grade 8 class with 23 students in low-income urban school in mid-western USA.</td>
</tr>
<tr>
<td>Tunstall (2001)</td>
<td>To analyze the social construction of the child, consequent power relations, and function of metaphors within classroom assessment discourse, and to consider these in relation to learning theory and policy contestation.</td>
<td>Qualitative: Critical discourse analysis using texts from an earlier (grounded theory) study on teacher feedback in elementary classrooms.</td>
<td>8 elementary teachers in 6 schools in London, UK.</td>
</tr>
<tr>
<td>Waddle &amp; Conway (2005)</td>
<td>To investigate the impact of a 5-year professional development partnership between a state university and a rural elementary school that aimed to improve student mathematics achievement.</td>
<td>Mixed: Explicitly mixed. Uses interview transcripts, grade-level meeting observation notes, classroom observation notes, teacher survey questionnaire, and student achievement scores.</td>
<td>450 elementary students, 19 teachers and 1 principal in Missouri, USA.</td>
</tr>
<tr>
<td>Yung (2001)</td>
<td>This study reports on the beliefs of teachers, and their influence on classroom assessment decisions and practices within a system-wide assessment scheme. Part of a larger project.</td>
<td>Qualitative: Interpretive narrative accounts drawing on taped lessons, observation notes, and interview transcripts.</td>
<td>3 experienced teachers in secondary science classrooms in 3 schools in Hong Kong.</td>
</tr>
</tbody>
</table>
at least one other source of influence sits between research and classroom assessment practices.

*Large-scale assessment.* The presence of large-scale assessment as a knowledge-generating source is not overwhelming in this study, with only nine related articles. Most frequently, large-scale assessment data is used to direct and justify educational change. Dori (2003) and Mason *et al.*, (2005) report large-scale results following an assessment reform programme, whereas Taylor and colleagues (2005), and Waddle and Conway (2005) use large-scale data to give credence to the success of professional development partnerships between schools and universities. Swanson and Stevenson (2002) use National Assessment of Education Programme (NAEP) data from across the United States to understand trends in standards-based reform. These studies all use large-scale assessment data to contribute to the knowledge base. Kyriakides (2002) takes the opposite angle in using research to recommend a large-scale assessment programme that collects baseline data for formative use in primary classrooms in Cyprus.

Only three articles look directly at the influence of large-scale assessment on teachers and their classroom practices. Boardman & Woodruff (2004) consider the influence negative, noting that state testing in Texas has become the reference point against which the utility of professional development and the feasibility of classroom change are measured by teachers. In Washington, teachers expressed the belief that large-scale testing has had a negative impact on classroom practices (Mabry *et al.*, 2003). Teachers and administrators surveyed in Alberta indicated a high degree of uncertainty about the impact of provincial examinations (McDonald, 2002), but a few interviewees felt that classroom assessment methods had improved. The discrepancy here may relate to contextual differences, or may simply be indicative of the range of opinions about assessment amongst educators. Of particular interest in McDonald’s (2002) study are responses relating to assessment literacy, which lead to the conclusion that educators ‘lack the professional training necessary to make the most of the available student-assessment information’ (p. 31). Low levels of assessment literacy in general may underlie the weak link seen here between large-scale assessment, as a source of information for change, and classroom practice.

*Evaluative inquiry.* Of the three-knowledge generating sources, evaluative inquiry plays the smallest role in this study. Only two articles relate to evaluative inquiry as it is defined in the framework. Dori (2003) discusses part of the evaluation of the Matriculation 2000 project, which is an assessment programme sponsored by the Israeli government, and Hayward *et al.* (2004) draw on the evaluative reports produced by the schools participating in the *Assessment Is for Learning* programme in Scotland. Evaluative reports are also used as a source of information for research by Mason and colleagues (2005), but the distinction between evaluative inquiry and research is blurred in their description of the reform efforts in a Californian school district. The authors identify their work as research, but the word ‘evaluation’ (p. 2)
appears in the purpose statement, and details are given concerning the merit, worth, and significance of the programme. Mason et al. (2005) produce, in effect, a hybrid of evaluative inquiry and educational research. A fourth article can be associated with evaluative inquiry, but through the results rather than the purpose of the research. Cheung (2002) finds evidence of informal programme evaluation by teachers in response to the expansion of a classroom assessment programme, the Teacher Assessment Scheme (TAS), in Hong Kong. The vast majority of the statements made by teachers who had yet to experience the TAS programme were evaluative in nature, and Cheung (2002) concludes that this type of evaluation plays a significant role in whether large-scale reforms are accepted or rejected by teachers in practice. A striking similarity in these four studies, beside their association with evaluation, is that they all involve large-scale efforts to change classroom assessment practices.

Mediating sources of influence

Educational policy. Although educational policy is mentioned in most of the articles in this study (17 of 24), it is the primary research topic of only one. To understand the effects of large-scale policy reform, Swanson and Stevenson (2002) examine 22 policies in 50 American states in relation to information about classroom practices from the NAEP teacher questionnaires, finding that ‘teacher receptivity factors’ (p. 15) strongly influence policy reform. The directness of the link between policy and teaching practice is reflected further in the current tendency to involve teachers significantly in educational reform (Dori, 2003; Hayward et al., 2004; Mason et al., 2005).

The indirect influence of policy on classroom practice is also evidenced in several studies. Often policy is named as the impetus for action, especially in relation to large-scale assessment and professional development. For example, three studies suggest that policy-makers’ responses to accountability pressures have led to an increase in large-scale assessment. Mabry et al. (2003) and Boardman and Woodruff (2004) point to the national No Child Left Behind policy in the United States, and McDonald (2002) identifies provincial examination policy in Alberta, Canada. Policy is also implicated by some as a negative influence on teacher education, and national policies are associated with concerns about professional development (Briscoe & Wells, 2002; Hall & Harding, 2002; Lovett & Gilmore, 2003). In one case, the impact of national policies on school priorities was considered detrimental to teachers’ assessment literacy (Hall & Harding, 2002).

While the effects of policy are repeatedly referred to in this collection of articles, there is also some evidence that central policies are mediated through local administration. Hall and Harding (2002) note the important role of the head teacher in implementing policy in England. Local policy administrators, such as head teachers or school principals, can facilitate change efforts through the allocation of time and funds for professional development and team meetings. Taylor et al. (2005) found that ‘high-reform-effort’ (p. 58) schools in five American states all benefited from the structural support provided by the school administrators (e.g., time in school schedules for group meetings), whereas schools that were less successful in changing
teaching and assessment practices received less administrative support. The interplay between central and local forces suggests that policy, in addition to being a source, plays a role in the stream between other sources of influence. Several authors comment on this dynamic, and some explicitly wish to inform educational policy. Kyriakides (2002) begins with this intent in developing a research-based model for assessment policy, whereas Cheung (2002) comes to the conclusion, after listening to involved teachers, that the Hong Kong Education Authority (HKEA) officers have much to gain in adapting assessment policy to respond to teachers’ concerns.

Professional development. Many of the teachers whose voices are heard through these studies have concerns about their preparation in the area of assessment. Teachers in Alberta (McDonald, 2002), Ontario (Hargreaves et al., 2002), and Washington (Mabry et al., 2003) all express frustration or concern about their lack of assessment expertise. The visibility of professional development in this group of articles (13 of 24) indicates that teachers are not alone in their concern. Several authors look at different forms of professional development in their quest to understand the learning process for in-service teachers. Lovett and Gilmore (2003) assist a diverse group of eight teachers using a ‘quality learning circle’ (p. 195) model in New Zealand, and Hand and Prain (2002) engage in a ‘reciprocal apprenticeship’ (p. 751) with two secondary science teachers during a four-year project in Australia. In the later study, and in another university-school partnership in Colorado (Borko et al., 2000), changes in the teachers’ assessment practices are traced over time. Taylor et al. (2005) implemented a more structured, university-based form of professional development with 93 teachers in 15 elementary schools across the United States. A number of factors that facilitate or hinder the change process, such as school environment and teacher characteristics, are identified in these studies.

Collaboration between university researchers and practitioners is identified in two studies as a form of action research that blends professional development and educational research (Torrance & Pryor, 2001; Briscoe & Wells, 2002). The teachers involved in these studies describe significant changes in their professional beliefs and assessment practices as a result of their participation, and the potential of professional development as a source of influence is highlighted. When members of a professional development project are like-minded, the change process is quick and easy. For example, Treagust et al. (2001) describe a collaborative team in the United States that was able to implement and document the successful use of embedded assessment over a brief, three-week period. However, most of the studies in this group strongly suggest that the flow from professional development to classroom practice is usually less direct and immediate, diffusing into multiple branches in the encounter with teachers’ existing beliefs.

Teachers’ beliefs and conceptions. Teachers’ beliefs and conceptions are not seen as the starting or ending point in any of these studies, but they are considered as part of the process in a significant number (14 of 24). Teacher’s approaches to learning were
mentioned in many of the studies involving professional development, and the passivity of some teachers as learners is particularly noted (Lovett & Gilmore, 2003; Hayward et al., 2004). Although these teachers were not resistant to innovation, they expected traditional instruction and were initially not prepared to direct their own learning. This reflects a teacher-centred approach to learning, whereas current assessment ideals are student-centred and consistent with the tenets of constructivism (Shepard, 2000; Roos & Hamilton, 2005). For the teachers in these studies, changing assessment practices entailed a shift in beliefs about teaching and learning. During a five-year collaborative professional development programme in Missouri, teachers expressed a decrease in beliefs about control as student-centred practices increased (Waddle & Conway, 2005). Briscoe and Wells (2002) found that critical reflection about beliefs was a precursor to changing assessment practices, but others note that the change process varies with individuals. For example, the two teachers in the Borko et al. (2000) study engaged in the change process differently as a result of their personal characteristics. In some cases, teachers’ beliefs are resilient, especially if they are supported by contextual or cultural factors, to the degree that they hinder positive change that might otherwise occur through professional development (Lock & Munby, 2000; Yung, 2001; Brighton, 2003; Boardman & Woodruff, 2004).

Teachers’ beliefs have the power to affect the implementation of policy. As noted earlier, the relationship between policy and practice is not always direct. Two studies that look at the role of teachers in policy implementation, Yung (2001) and Cheung (2002) are interestingly both concerned with the Teacher Assessment Scheme (TAS) in Hong Kong. Although they differ considerably in their methodology, they come to a common conclusion. Both state that TAS policies must address teachers’ beliefs if the programme is to be fully adopted, rather than partially adapted, by classroom teachers. In her analysis of assessment discourse, Tunstall (2001) suggests that when teachers’ beliefs are in conflict with policy, such as the National Curriculum Assessment policy in England, teachers may have the final say:

Practitioners do not confront policy texts as naive readers; they come with their own histories, purposes, values and vested interests in the meaning of policy...while policy makers may attempt to control for correct meaning, their texts enter existing power relations in terms of discourse and they rely on key mediators for success. (p. 229)

While teachers’ roles in mediating assessment reform are emphasized in several studies (Tunstall, 2001; Swanson & Stevenson, 2002; Hayward et al., 2004), there is also considerable evidence that teachers are influenced in the reform process. Teachers that participated in the trial of a new approach to secondary science assessment experienced a ‘conceptual change’ (Dori, 2003, p. 39) that affected their practices. Cheung (2002) correlates changes in teachers’ ideas to their length of experience with the TAS in Hong Kong. Teachers with no TAS experience were sceptical of its value, and teachers with some experience expressed practical concerns about workload and procedures. The most experienced teachers had shifted to more complex concerns about the consequences of student assessment, and the lack of collaboration amongst teachers, and they were also considering how TAS could be improved. Changes in
teachers’ ideas over time are also documented by Mabry et al. (2003) in relation to standards reform, and they are compared by McDonald (2002) to earlier perceptions about provincial testing.

Teachers’ understandings about assessment are central to the change process considered in this paper. In several cases, teachers are studied while they construct knowledge about student learning. Tunstall (2001) looks at teachers’ explicit assessment comments, whereas Morgan and Watson (2002) attempt to examine the internal decision-making by teachers in two separate studies in England. In the first, a teacher is observed directly during classroom assessment activities, and in the second a group of teachers engage in a think-aloud process while evaluating passages of student writing. According to the researchers, the numerous factors involved in the assessment, including the variety of resources used by the teachers, the characteristics of individual students, and the ongoing reform context, introduced considerable inequity in the process. Morgan and Watson (2002) conclude with a call for the ‘raised awareness of processes of formative assessment’ (p. 104), and this push for assessment that supports student learning is at the heart of much of the human activity displayed in these articles.

The feedback arrow: practice as a source of influence

Assessment practices in elementary and secondary classroom contexts were situated as the end-game in the conceptual model described earlier, but a feedback arrow to the knowledge-generating sources was included to indicate that practice may also be the basis for knowledge generation. Evidence of this relationship was not overwhelming in the sample, but four studies did clearly make the link back to practice. The most obvious connection is made by Torrance and Pryor (2001), who redesign their model of formative assessment after working collaboratively with classroom teachers in England. Briscoe and Wells (2002) team up to ground research on portfolios in the logistical realities of an elementary classroom in Florida. Two studies on professional development also use practice to reconceptualize theory (Lovett & Gilmore, 2003), or more simply, to learn from teachers (Hand & Prain, 2002). These works all involve collaboration between researchers and practitioners, and they function, in effect, to close the loop in the information process.

Discussion

Emanating from Australia, Cyprus, Hong Kong, Israel, New Zealand, England and Scotland in the United Kingdom, the provinces of Alberta and Ontario in Canada, and Missouri, Texas, Washington, and Virginia in the United States of America, the international nature of the research on classroom assessment is clearly seen in this selection of articles. Given the contextual nature of education, differences arise from their geographic diversity. However, many commonalities also surfaced during the analysis. In the thematic discussion that follows, three cross-study currents, and two significant tensions are considered.
Common currents

Research perspective. A noticeable trend relating to research perspective is easily discernable in the methodology and the underlying epistemology of the studies. Of the 24 articles analysed, 23 make use of some form of qualitative data, with 16 of these relying on observation, interviews, field notes, and documentary evidence. In other words, the majority of the articles follow a qualitative tradition, with a smaller group using mixed methods (7 of 24), and only one based solely on quantitative data. This methodological homogeneity is reflected in a similar, but weaker epistemological trend. Although an epistemological stance is not explicitly stated in a large number of the studies (14 of 24), a theoretical orientation is often indicated that places most within an interpretive framework. Of the studies in which the epistemology is specifically identified, six approach research from a constructivist perspective (Borko et al., 2000; Torrance & Pryor, 2001; Briscoe & Wells, 2002; Hand & Prain, 2002; Morgan & Watson, 2002; Mabry et al., 2003), one adopts a critical poststructuralist perspective (Tunstall, 2001), and a final article makes use of multiple perspectives (Hargreaves et al., 2002). This trend is interesting for several reasons. First, the skewed nature of the sample was not an expected result given that methodology was not considered in any way during the selection process. Second, the dominance of qualitative methods in studies on assessment, traditionally a quantitative field, suggests that shifting beliefs may be attributed to researchers as well as teachers. These studies are all concerned with processes that involve a high degree of human activity, and in many cases, their dynamic nature is acknowledged through the choice of methodology.

Collaboration. A second notable current is the theoretical commensurability between constructive assessment, and the social approach toward the construction of knowledge that these studies describe. Collaboration appears in a variety of shapes and forms. There is collaboration in the research process (8 of 24), in professional development activities (9 of 24), and even in large-scale reforms (3 of 24). One research team goes so far as to use the words collaboration, collaborate and collaborative over 35 times in a ten-page text (Lock & Munby, 2000). Many of the collaborative undertakings that are described in these studies forge a strong link between the research and teaching communities through university/school partnerships or action research projects (Borko et al., 2000; Torrance & Pryor, 2001; Treagust et al., 2001; Briscoe & Wells, 2002; Hand & Prain, 2002; Hayward et al., 2004; Waddle & Conway, 2005). This emphasis on collaboration is consistent with recommendations in the literature that discusses educational change in more general terms (Horn, 2002; Fullan, 2003), and collaboration can be considered a hallmark of the relationship between educational research and professional development, as two sources of influence operating in these articles.

Seen together, these works also highlight areas of limited collaboration. Calls for more effective use of existing knowledge have recently been heard in research circles (Gorard, 2002; Levin, 2003), and information that is generated through evaluative
Influences on classroom assessment

Time: a frequent factor. Time is mentioned frequently in the analysed articles, and it is considered a third strong current. When assessment reform is undertaken on a large-scale, the political will to work with long policy time-frames is a prerequisite (Hayward et al., 2004), and evidence of the effectiveness of an innovation should be gathered over multiple years (Mason et al., 2005). Taylor et al. (2005) note that earlier measurements may not accurately reflect the effect of reform, and they conclude that the change process needs to be supported for a sustained period for teachers to improve their assessment practices. Local structures and administrators should also support teachers’ constructive assessment practices, and this can be achieved through the allocation of time for professional activities, collaborative research, and common schedule blocks. However strongly recommended, this is still not the reality experienced by many teachers. Overloaded teaching schedules affect participation in research projects (Lock & Munby, 2000; Torrance & Pryor, 2001; Brighton, 2003), and assessment activities that are not conventionally associated with instruction, such as sharing assessment data or discussing practices with colleagues, are usually done in personal time (Hall & Harding, 2002; Ingram et al., 2004). The imposition of additional professional activities, such as evaluative inquiry or assessment data analysis, without the allocation of time overloads teachers (Cousins, 1998; Earl & Fullan, 2003), which in turn reduces their ability to engage in creating change (Lovett & Gilmore, 2003; Ingram et al., 2004).

A shortage of time is often given as a reason for maintaining traditional assessment practices in the classroom. In the early stages of the change process, teachers worry that constructive assessment practices will take time away from ‘covering the curriculum’ (Borko et al., 2000, p. 291), and they believe that new assessments will be too time-consuming (Torrance & Pryor, 2001; Cheung, 2002; Hargreaves et al., 2002; Mabry et al., 2003). Even those who appreciate the potential of assessment for learning report that it demands more time in practice (Morgan & Watson, 2002; Dori, 2003). The perceived shortage of time in education has been discussed repeatedly (e.g., Hargreaves et al., 2002; Briscoe & Wells, 2002; Fullan, 2003), and classroom assessment researchers continue to identify time as a significant issue for practitioners (Gilsdorf, 2000; Schmidt, 2001). In this study, Briscoe and Wells (2002) are unique in offering a positive means of dealing with time constraints. They describe how one teacher saved classroom time by using assessment formatively. As time has clearly been identified as a hindrance factor in the implementation of classroom assessment
that supports student learning, research that sheds light on ways to reduce time issues would certainly be beneficial for all stakeholders.

*Tensions in the change process*

In considering how sources of information influence classroom assessment practices, two tensions became apparent. The first relates to the traditions of teacher autonomy and school community, and the second involves the relationship between collective commitment and the development of assessment literacy.

**Balancing autonomy and community.** Changing assessment practices is not a simple process for teachers, and some feel it is more difficult than changing other teaching practices (Borko *et al.*, 2000; Earl & Katz, 2000). Even some of the motivated teacher-volunteers in these studies faced difficulties (Lock & Munby, 2000; Lovett & Gilmore, 2003), and those who were successful in changing their assessment practices mention supportive partners and leaders (Briscoe & Wells, 2002; Waddle & Conway, 2005). Cousins and Leithwood (1993) suggest that these types of relationships are critical in that ‘opportunities for social interaction must be available for organizational learning to take place’ (p. 330). Benefits are clearly seen from a collaborative environment with strong, supportive leadership (Hall & Harding, 2002; Hayward *et al.*, 2004; Taylor *et al.*, 2005), and schools that lack harmony face considerable challenge and frustration in the change process (Cousins, 1998; Datnow, 2000). The role of school culture is evident in the literature on school improvement (Datnow, 2000; Sutherland, 2004), and it is also emphasized in research on classroom assessment (Stephens, *et al.*, 1995; Delandshere & Jones, 1999; Stiggins, 2001; Carless, 2005). In the studies analysed here, several authors note difficulties in changing assessment practices when traditional beliefs about teaching, learning, and assessment dominate the learning community or culture (Lock & Munby, 2000; Brighton, 2003; Boardman & Woodruff, 2004). In this light, the community may facilitate or hinder the change process.

Teachers tend to be sceptical about mandated change (Bailey, 2000), and in many contexts they have enjoyed a long history of autonomy (Cuban, 1990). Some teachers believe that large-scale reform, especially if it involves standardized testing, undermines their professional judgement (Mabry *et al.*, 2003) and reduces their autonomy (McDonald, 2002). The importance of teacher involvement in the reform process is discussed in the literature on educational change (Cousins & Leithwood, 1993; Horn, 2002; Fullan, 2003; Sutherland, 2004), and it is further evidenced by the benefits arising from the active and meaningful participation of the teachers in these studies (Torrance & Pryor, 2001; Briscoe & Wells, 2002; Hand & Prain, 2002; Mason *et al.*, 2005). While purely bottom-up strategies are not the answer (Horn, 2002; Fullan, 2003), it is clear that teacher participation in the change process is essential. As such, teacher agency, rather than autonomy, is needed to support the move toward more constructive classroom assessment practices.
Quick fix or comprehensive commitment. The process of changing assessment practices is neither direct, nor immediate. Much of the work considered in this paper involves long-term commitment. Several policy-oriented studies consider the impact of programmes in their fourth or fifth year (Dori, 2003; Mason et al., 2005), and the studies that focus on professional development involve projects that were sustained over two to five years (Borko et al., 2000; Hand & Prain, 2002; Brighton, 2003; Taylor et al., 2005; Waddle & Conway, 2005). Some of the researchers also engage in lengthy data collection periods (Hall & Harding, 2002; Kyriakides, 2002), and others examine data longitudinally (McDonald, 2002; Swanson & Stevenson, 2002). According to Fullan (2003), expecting a slow pace of change is the first lesson in understanding complex educational reform.

The complexity of change is highlighted in several studies where new assessment methods lead to broader changes in teaching practices, which ultimately alter the learning environment (Treagust et al., 2001; Dori, 2003; Waddle & Conway, 2005). Changing assessment practices is not simply a matter of increasing teachers’ assessment literacy through professional development workshops, but a more comprehensive process that requires a conceptual shift for all stakeholders. Conflicting ideological messages about teaching, learning, and assessment are often found in educational policy and procedures (Delandshere & Jones, 1999; Earl & Katz, 2000; Tunstall, 2001; Hargreaves et al., 2002; Hayward et al., 2004). Not only does the impact of external messages on classroom assessment need to be considered, but the influence of teachers in how reform efforts are adopted, adapted, or rejected must also be taken into account.

The use of information by teachers is a key factor in the move toward assessment practices that support student learning. Mabry et al. (2003) found that some teachers’ hold a low opinion of university-based training, and identify their colleagues as the most important source of information, which is consistent with the results of broader surveys (Landrum et al., 2002; OCT, 2003). When considered in tandem with the professional development experiences described by the teachers in these studies, the preference for information that is immediately and contextually relevant suggests that collaboration should be a feature of the ongoing effort to improve assessment practices. Rather than one-shot training sessions delivered by assessment experts, more collaborative professional development projects that significantly involve educators at the local level will be needed.

Conclusion

In summary, all the sources included in the conceptual model for this paper do play a role in the dynamics that are evidenced by the articles. Of the three knowledge-generating sources, the influence of educational research is most evident, but it is clear that the universal reliance on research as a starting point in the change process does not lead directly to changes in classroom assessment practices. Most of the activity seen in the articles involves the mediating sources of influence. The frequency with which they are discussed is fairly evenly distributed between the two external mediating
sources (educational policy and professional development), and the internal mediating source (teachers’ beliefs and conceptions). The relationship between the sources of influence can be described as dialectic in nature, as influence often flows between them and affects their initial form. However, the dynamics seen in these articles are far too diverse to be modelled as a single dialectic, and the complex patterns that emerge can best be described as a postmodern dialectic. Unlike the duality and symmetry of the traditional Hegelian dialectic, the postmodern dialectic models tension between multiple and unequal perspectives that are mediated by a variety of stakeholders to form an ongoing series of dynamic scalene triangles. It is through these interactions that new knowledge about classroom assessment is created, negotiated, and practised. An implication for both research and practice illuminated by these studies is that there is no single right way to understand the change process, and no universal solution for changing classroom assessment practices.

This research brings together a group of studies on a variety of organizational levels that are informed by a range of perspectives. Collectively, the articles give credence to considerable movement in the area of classroom assessment, and they mitigate the evidence of inertia that introduced this paper. Although they may involve small pockets of exemplary teachers, these studies do indicate that there is a considerable effort to effect change in classroom assessment practice through the implementation of policy-based or professional development programmes around the world. This study highlights the critical need for sustained and cooperative effort in changing classroom assessment practices. Researchers, policy-makers, administrators, and teachers, whose needs and perspectives have frequently differed, are now being asked to work together in this change process (Hayward & Hedge, 2005; Invernizzi et al., 2005). Sources that are currently not well represented in educational literature, such as evaluative inquiry and large-scale assessment, could be used more effectively to construct a better understanding of the issues involved. The combined influence of knowledge-generating and mediating sources will be necessary to encourage assessment practices that support student learning, and ultimately permit classroom assessment to fulfil its pedagogical potential.

Notes on contributor

Robin Tierney is a Ph.D. Candidate at the Faculty of Education, University of Ottawa in Ontario, Canada. A former elementary teacher and reading specialist, her research focuses on the use of fair assessment to support student learning. Robin Tierney is a member of the Measurement, Evaluation, and Assessment Research Unit (MEA) at the University of Ottawa.

References


Influences on classroom assessment


