Towards Balanced Assessment of Student Teaching Performance

Keith Roscoe

University of Lethbridge
Towards Balanced Assessment of Student Teaching Performance

Keith Roscoe
University of Lethbridge

Abstract

Assessment practices in schools have undergone dramatic changes over the last decade, and applying this knowledge to the assessment of student teachers is a challenge currently facing teacher preparation programs. K-12 assessment has moved towards a "backwards design" approach, greater student involvement, a wider range of strategies, and assessment systems that balance summative and formative assessment. However, the assessment of student teaching performance during field experiences has often overemphasized summative assessment, collecting data for making judgments, at the expense of formative assessment, gathering information to improve student teacher performance. Recently, one institution recognized the need to reexamine its approach to field experience assessment based on the thrust towards 21st century education, the growing knowledge base in assessment, and feedback from its educational partners. The article is a case study of this improvement initiative: the context, process involved, the outcomes of the improvement process, and implications for teacher education.

The basic premise of the current impetus towards 21st century learning is that the world has undergone so fundamental a change over the last few decades that what people learn and how they learn it has likely changed forever (Trilling & Fadel, 2009). Schools must be redesigned to meet present and future needs of students, with students working collaboratively in teams on learning projects, solving real-world problems, and building their critical thinking, communication, and innovation skills. A complex 21st century curriculum with a heavy focus on skills calls for movement towards a more comprehensive and balanced approach to assessment at all levels of education (Partnership for 21st Century Skills, 2009, 2010). Teacher preparation programs will play a key role in the necessary educational reforms, and assessment strategies for 21st century knowledge and skills will be a crucial component of not only the curriculum, but also the assessment practices of teacher education (Partnership for 21st Century Skills, 2009, 2010).

At the same time that momentous changes in technology, the economy, and the workplace are driving the 21st century skills movement, great strides have been made in our understanding of effective teaching and learning, and a consensus has emerged around professional standards for teachers that reflect this knowledge base (Darling-Hammond & Bransford, 2005). Assessment practices in schools, for instance, have undergone dramatic changes over the last decade. The traditional approach to assessment typically involved a narrow range of assessments focused on collecting information on student achievement for grading purposes, often misaligned with curriculum outcomes, with a mainly passive role for students. The current understanding of classroom assessment emphasizes balanced assessment—using a variety of assessment practices to both measure learning and foster gains in learning, while tightly linking assessment to outcomes, and actively engaging students in assessment (Black & Wiliam, 1998a, 1998b; Black, Harrison, Lee, Marshal & Wiliam, 2003; 2003; Davies, 2007; Stiggins, 2000; Stiggins, Arter, Chappuis & Chappuis, 2006; Sutton, 1995). Although teacher
education programs have, in many cases, adjusted the curriculum of teacher education and embraced professional standards to reflect this approach to assessment, the actual practices used to assess student teacher performance have often lagged behind. This was the situation in a Western Canadian teacher preparation program a few years ago. Spurred on by new provincial assessment standards for student teachers and partnerships with professional organizations, there was a recognized need to improve the assessment of student teachers. A team of faculty members was tasked with creating a better alignment between assessment in the field experiences component of the program and what student teachers were learning about assessment in their courses. This article is a case study of this improvement initiative: the context, process involved, the outcomes of the improvement process, and some implications for teacher education.

Changing Ideas and Practices in Assessment

Assessment in K-12 Education

Over the last decade or so, there has been a major re-thinking of educational assessment and evaluation in North America and elsewhere (Black & Wiliam, 1998a, 1998b; Black et al., 2003; Davies, 2007; Stiggins, 2000; Stiggins et al., 2006; Sutton, 1995). A changing society has placed greater demands on teachers, with more diversified classrooms and broader, more complex 21st century curricula in a range of school subjects. In the meantime, research has provided a more complex view of learners and learning and a wider range of well-tested teaching and assessment strategies. The aims of assessment have greatly expanded from simply evaluating work and assigning grades to a much wider range of purposes, such that the overall thrust is away from determining and reporting grades and towards improving learning by integrating assessment with everyday planning and instruction. Furthermore, the work of Wiggins and McTighe (2005) on understanding by design or "backwards design" has had considerable influence on educators' efforts to align assessment, planning, and instruction with educational outcomes.

In Western Canada, and Alberta in particular, this shift in ideas has had a significant impact on policy and practice in school assessment. Provincial teaching standards in classroom assessment for in-service teachers (Alberta Education, 1997), detailed standards for pre-service and beginning teachers (Alberta Education, 2006), and a recent report outlining a long-term vision for Alberta's 21st century education system (Alberta Education, 2010), exemplify the call for balanced assessment, student involvement, more valid assessment tools and techniques, and effective communication and reporting of assessment results. In addition, the Alberta Initiative for School Improvement (Alberta Education, 2010), a provincial government initiative which funds projects to improve student learning, has had a strong focus on improving classroom assessment over the last several years. Overall, it is fair to say that there is a strong emphasis on improving K-12 assessment practice in the province of Alberta.

Assessment in Teacher Education

In North American teacher education, it seems that more attention has been paid to what teachers need to learn about assessment than to rethinking how student teachers are actually assessed. In the U.S., for instance, the NCATE Standards for Student Teaching (National Council for Accreditation of Teacher Education, 2008), call for decisions about student teacher
performance to be based on multiple assessments made at several points, and for the establishment of fair, accurate, and consistent assessment procedures. The National Academy of Education Committee on Teacher Education (Darling-Hammond & Baratz-Snowden, 2005) suggests that successful student teaching requires clear standards for performance, adequate opportunities for practice with continuous formative feedback and coaching, and structured opportunities to reflect upon and improve practice. Bransford, Darling-Hammond, and LePage, (2005), and Sheperd, Hammerness, Darling-Hammond, and Rust, (2005) advocate that beginning teachers learn how to use standards to construct assessments, employ a large repertoire of formative assessment strategies, and help students learn to self-assess. However, there is relatively little focus in these documents on improving assessment of student teachers beyond expanding the range of assessments to include performance assessments, and placing a greater emphasis on self-assessment using portfolios.

Similarly, in the report of the AERA Panel on Research and Teacher Education, discussions of related research focus on the structure and outcomes of teacher preparation programs (Zeichner & Conklin, 2005) and the impact of methods courses and field experiences on student teachers’ educational beliefs and practices (Clift & Brady, 2005). Touching more directly on the assessment of student teacher performance, Castle and Arends (2006) and Arends (2006) concentrate on the development of effective performance assessments. The authors do describe assessment of teaching performance (e.g., observations, videos, and microteaching), but most of the performance assessments involved course-based tasks such as lesson plans, unit plans, case studies, projects, portfolios, and reflective journals. In all, there appears to have been relatively little attention paid to systematically applying the research on K-12 assessment to the assessment of student teachers’ teaching performance.

A Case Study: Improving Assessment in Teacher Education

The Teacher Preparation Program

The university is a mainly undergraduate institution with a student population of around 8000. The teacher preparation program is offered either as part of an undergraduate combined degree program, or as an after-degree program, with about half of the 216 students enrolled annually pursuing each option. The combined degree program is a five-year 50-course program in which students complete at least 30 courses in a bachelor’s degree in their major and 20 semester courses in a bachelor of education. Students with a completed degree complete a four-semester, two-year program. Before being admitted, all applicants must successfully complete an Orientation to Teaching course which provides an opportunity for mentor teachers and faculty to assess candidates’ teaching potential and for students to spend substantial time in the classroom and find out more about the teaching profession and its challenges.

The program places a high priority on field experiences, in that each student accumulates 28 weeks of student teaching, and faculty are heavily involved in practicum supervision as university mentors. During their program, education students take three professional semesters. The first professional semester is focused on basic teaching skills across all subject majors, and requires a six-week teaching practicum at the elementary or middle school level. In the second professional semester the focus in both coursework and the six-week practicum is on the student’s teaching major. Finally, in the third professional semester, or internship, students are assigned to a school full-time for 15-16 weeks with a 50% teaching: 50% professional
development assignment. Graduates of the program are eligible for certification to teach from Kindergarten to Grade 12 in the Province of Alberta. It should be noted that there are no national standards or accreditation process for teacher education in Canada along the lines of the American NCATE standards (National Council for Accreditation of Teacher Education, 2008), and teaching standards are developed at the provincial level. In Alberta, the Teaching Quality Standard (Alberta Education, 1997) sets out the standards for graduates of teacher preparation programs as well as for in-service teachers.

The Need for Improvement

Prior to 2005, the assessment system for field experiences in the first two professional semesters consisted of the following elements: (a) daily observation of teaching performance followed by oral and written feedback from teacher mentors, (b) weekly observation of teaching performance followed by oral and written feedback from university mentors, (c) completion of a formative competency checklist by teacher mentors at intervals over the practicum, (d) the completion of a summative field experience report by the teacher mentor and university mentor at the end of the practicum. Informal comments from teacher mentors, principals, faculty mentors, and student teachers indicated that although the existing system was satisfactory, there was a need for a more valid and reliable method for gathering information and making accurate judgments about student teaching performance.

For one thing, the standards and criteria contained in the formative instrument did not accurately reflect current “good practice” in K-12 education or emerging 21st century skills as embodied in provincial teaching standards and the literature. Critical aspects of student teaching performance in areas such as classroom management and leadership skills, professional career skills, and ICT integration skills were poorly addressed or not addressed at all. Other key areas of performance such as planning and preparation, assessment, and instruction needed updating to reflect advances in the knowledge base in curriculum design, differentiation, and assessment as well as to bring the standards in line with current provincial standards and program outcomes (which had been strongly influenced by the 21st century approach). Another issue raised by stakeholders was the vague or ambiguous wording in some of the standards. Teacher mentors and faculty mentors reported that in some cases it was difficult to fail poorly performing students or to fully recognize and document outstanding teaching performances using the standards in the existing assessment instruments.

In addition, stakeholders indicated that the format and structure of the instruments for assessing teaching performance in the field needed adjusting. They pointed out that the existing formative instrument did not facilitate regular and effective feedback by mentors and peers or regular self-assessment by student teachers. The checklist scale for rating teaching performance, consisting of a continuum from Weak to Strong, was regarded by respondents as inadequate for diagnosis, feedback, and decision-making. In addition, stakeholders pointed out that the summative instrument did not list the standards being addressed or include a rating scale, but merely spaces for comments and for recording a grade of Pass, Fail or Incomplete.

The Improvement Process

The process for improving the assessment of first professional semester student teaching performance was a fairly complex one, consisting of several stages: (1) consultations with
stakeholder groups, (2) research on relevant teaching performance standards and assessment design, (3) revision of standards, (4) redesign of assessment instruments and procedures, (5) feedback on new instruments and procedures from stakeholder groups, (6) piloting of new instruments and procedures, (7) further adjustments to instruments and procedures based on stakeholder feedback and the piloting process, (8) implementation of an improved assessment system for teaching performance. This process was carried out over a two-year period by a design team comprised of program administrators and the four professional semester coordinators, including the author, who, at that time, was the Teaching Internship (PSIII) Coordinator.

The consultations carried out before and after the re-design of the teacher performance assessments (stages 1 and 5), involved five focus groups (N = 5 to 20 participants) conducted with groups of faculty members, local and provincial representatives of the provincial teacher association, and regional school district administrators. The guide questions for the focus groups solicited feedback in three areas. They were: (a) effective and successful aspects of the present practicum assessment system (What do you like about the present system? What is working? What should we keep doing?), (b) ineffective aspects and areas for improvement (What don’t you like about the present system? What isn’t working? What should we stop doing?), and (c) concrete suggestions for improvement (Which standards should we add, remove or modify? How should the assessment instruments be structured and formatted? How can we improve assessment procedures in the practicum?).

The research and development phase of the process (second and third stages) was carried out by subcommittees of the team working on the various areas of teaching performance such as planning and preparation, instruction, classroom management, assessment, and professional standards. Besides the provincial teaching standards and input from stakeholders, major points of reference included Danielson’s teaching framework (1996), Wiggins and McTighe’s (1996) work on backwards design, Evertson and Emmer (2003) and Bennett and Smilanich (1994) on classroom management, and assessment ideas from the Alberta Assessment Consortium (2004), Black et al. (2003), and Stiggins (2000). After several months of vigorous debate and deliberation, a consensus was reached on design of the instrument and the standards to be used.

Following further consultations with faculty, teachers and administrators and additional revisions of the instruments and assessment procedures, the new assessment system for the Professional Semester One Practicum was field-tested in Fall 2005 with 180 mentor teachers and their student teachers. Mentor teachers were provided with the new instruments and a set of suggested guidelines for use before the beginning of the practicum (University of Lethbridge, 2007a). University mentors attended workshops explaining the new instruments, standards, performance levels, and guidelines for use, with a view to sharing this information with their assigned teacher mentors.

At the end of the PSI practicum, a user survey was used to solicit feedback from mentor teachers on the assessment instruments. Respondents (N = 116; return rate = 64%) strongly agreed or agreed that the new formative assessment form was clear and easy to understand (91%), easy to use (91%), and useful for providing feedback (87%). Many detailed and useful suggestions were received and incorporated into the next round of revisions (University of Lethbridge, 2006, 2008). One suggestion that was not implemented was to write detailed descriptors for the three performance levels (Not Meeting Expectations/Meeting Expectations/Exceeding Expectations) for each of the 62 standards in the formative and summative instruments. It was soon determined that neither a performance assessment rubric nor
a separate set of descriptors on this scale was feasible. The decision was made to address reliability issues by relying on the rating scale, supplemented by training sessions attended by university mentors and teacher mentors in which our common understandings of each performance level were explored. With very few exceptions, this has proved to be an acceptable resolution.

Meanwhile, the revised Professional Semester One (PSI) instruments and procedures were used as the basis for developing and piloting the new assessment practices in the Second Professional Semester (PSII) Practicum during March-April 2006. Consultations with faculty specialists in curriculum, psychology, educational foundations and assessment, as well as with teacher and administrators representatives, helped define the progression in standards from basic generalist teaching in PSI to an intermediate subject-specific focus in PSII (University of Lethbridge, 2007b, 2007c). Otherwise, the format and structure of the instruments, and the Guidelines for Use were retained (University of Lethbridge, 2007a). Once again, there was a training session for university mentors, and mentor teachers were asked to provide feedback on the new instruments and procedures by means of a paper and pencil survey. Feedback from PSII mentor teachers from a user survey of the new form was also very positive. For instance, teachers surveyed (N = 61) strongly agreed or agreed that the form was easy to use (93%), appropriate for the level (97%), and made it easier to evaluate students (95%). During the next academic year (2007-2008), the new assessment system was fully implemented in professional semesters one and two.

**Improvement Process Outcomes**

**Alignment of assessment system components.** Figure 1 shows how, in keeping with a backwards design approach (Wiggins & McTighe, 2005), the whole assessment system for field experiences is built around the revised criteria and standards for student teaching performance. The assessment tools and techniques are derived from, and consistent with, the desired standards, and are designed to gather good evidence of student achievement of the standards. Explicit communication about teaching performance standards by way of mentors and assessment instruments provides student teachers with clearer learning targets. Both formative and summative tools and procedures are designed to provide users with accurate information about learning in reference to the standards, whether for helping students diagnose and improve their performance, or for helping mentors make judgments about student teacher performance. Suggested tools and procedures for student self-reflection encourage student teachers to examine areas of strength, areas for improvement, and improvement goals in relation to standards for student teaching performance. Guided by the same standards, mentors can use assessment evidence and results to modify assessment tools and techniques, guide professional learning, and improve supervision practice.

**Balanced assessment.** In a balanced assessment system, formative assessment and summative assessment work together to produce an outcome that is greater than the sum of its parts (Stiggins et al., 2006, Chappuis & Stiggins, 2008). Purposefully planned summative and formative assessments combine to create a more accurate and complete picture of student learning. Some progress towards a more balanced system for assessment of student teaching performance has been achieved in the program. Feedback from teacher mentors, school
Figure 1. A Balanced Assessment System for Field Experiences

administrators, and university mentors suggests that improvements in the summative instruments and procedures have enhanced our ability to accurately measure and report teaching performance at the culmination of practicum experiences. Summative assessment has been balanced with better formative instruments and processes and enhanced opportunities to provide student teachers with timely useful feedback to help them improve their performance and meet the standards.

The formative assessment instruments were designed to help teacher mentors provide student teachers with specific written feedback relative to performance standards at regular intervals during the practicum. To enhance the quality and quantity of feedback, mentors are urged to concentrate on the three to five most relevant and significant strengths and areas for growth when they provide written feedback using the form, and to offer feedback in their observation notes that is closely linked to the standards described in the assessment forms. However, the process of balancing the formative and summative functions of assessment in the practicum has just begun. More remains to be done to articulate the summative and formative components of the system, to help students fully understand the intended standards and criteria, and to provide accurate and descriptive feedback.

Student involvement in assessment. A key to quality assessment at any level is the active involvement of students in the assessment process, such as teaching students to self-assess and to set goals (Stiggins et al., 2006). A self-reflection tool was developed to guide student teachers in the assessment of their own performance daily, encouraging them to take an active role in post-observation conferences with mentors. The self-reflection tool asks student to (a) describe the actual lesson, comparing what happened during the lesson to the lesson plan and
characterizing students’ responses to the lesson; (b) identify the main strengths (e.g., 3-5 strengths) and weaknesses (e.g., 1-2 weaknesses) of the lesson plan and the lesson-as-taught, (c) identify improvements to the lesson plan and the instruction of the lesson, and/or alternative ways of teaching the lesson; and (d) reflect on their professional growth in relation to the lesson, their professional learning goals, and their vision of teaching as described in their professional growth plan or teaching portfolio.

Student teachers develop a professional growth plan and/or portfolio for the practicum in which they set goals for professional learning based on the teaching performance standards. The professional growth plan and portfolio engage students in self-reflection, and allow them to document their professional learning and share evidence of achievement at conferences with teacher mentor/university mentors. In addition, student teachers placed in the same school or neighboring schools are encouraged to visit each other’s classrooms and offer specific constructive feedback, using the same formative forms used by teacher mentors. These initiatives to increase active student involvement in assessment in the practicum have had considerable success, but have been limited in some cases by a lack of consistency in application, and by student teachers’ and teacher/university mentors’ greater experience with and predisposition towards a more passive learning role.

Conclusion

Educators have come to understand that students learn best when teachers use a wide range of assessment methods to monitor learning, provide students with timely feedback, help students understand what is expected, and fully involve them in the assessment process (Black & Wiliam, 1998a, 1998b). As well, there is a pressing need for effective assessments focused on 21st century knowledge and skills (Partnership for 21st Century Skills, 2010). Applying these understandings to teacher preparation is a major challenge facing teacher education today. In this article, I set out to document how one institution approached this challenge, as they tried to align practicum assessment with research-based assessment practices and emerging professional standards, while meeting the practical needs of various stakeholders.

In terms of the 21st century approach, the program was able to incorporate several key 21st century program components recommended by the Partnership for 21st Century Skills (2010). A leadership team set direction for changes in the assessment of clinical experiences and engaged faculty in developing performance-based assessment built on professional standards. The assessment redesign process involved partnerships with K-12 schools, and local and provincial professional organizations and resulted in a more balanced assessment system for our already extensive and coherent field experiences. The extensive collaborations with partners in the educational community, involving consultation, research, development, and further consultation, were a key to developing assessment tools and techniques that met the varied expectations and needs of stakeholders. However, regarding education for the 21st century, this is just the beginning of needed changes in field experience assessment. Program standards, coursework and assessments will need to be updated to reflect knowledge and skills that are presently absent or underrepresented, for instance, global awareness, entrepreneurial literacy, creativity/innovation, information/media literacy, social/cross-cultural skills, interdisciplinary project learning, and greater levels of collaboration during field experiences (Alberta Education, 2010; Partnership for 21st century skills, 2009).
In terms of the call for movement towards balanced assessment systems (Partnership for 21st century skills, 2009), the results suggest that effective research-based K-12 assessment practices may also be effective in teacher education. Based on the feedback from stakeholders such as teachers, administrators, students, and university mentors, substantial improvements were achieved in providing a broader base of formative assessments to complement summative assessments, in involving students in the assessment process, and in tightening the coupling between standards, instruments and procedures. The experiences of this faculty of education do suggest, however, that applying the notion of balanced assessment systems to teacher education is not a simple straightforward task.

Necessary conditions for an innovation of this kind include adequate time and resources committed to the project over a period of years, a consensus within the program about the need for change and the nature of the changes, and effective leadership to see the improvements through to completion. Furthermore, this case does not necessarily generalize to other programs, given that this is a relatively small-scale, well-funded program, with heavy faculty involvement in field experiences, and support for innovation from within the institution and various partners with an interest in teacher education. However, given the potential of effective research-based classroom assessment for improving student learning, and the need for teacher education programs to model these practices, these findings may have some significance in illustrating the possibilities for change, and what a balanced assessment system might look like in the field component of teacher education.

References


Alberta Education. (2006). *Effective Student Assessment and Evaluation in the Classroom: Knowledge and Skills and Attributes*. Edmonton, AB: Alberta Education.


