

THE COST OF COMPLIANCE: UNIT STANDARDS FOR TEACHER EDUCATION IN THE LIGHT OF RECENT EXPERIENCE IN SECONDARY SCHOOLS

ALAN HALL¹ AND PAUL KEOWN²

1. *Department of Professional Studies
University of Waikato*
2. *Department of Social, Physical, & Health Education
University of Waikato*

ABSTRACT *Observations of the demands on the work of teachers created by a year long trial implementation of unit standards in secondary schools highlighted an initial tendency for the assessment requirements to dominate teaching and the need for teachers to develop new strategies and systems to combine assessment with teaching, track the achievements of individual students and cope with the practical requirements of re-assessment. Towards the end of the trial some evidence was noted of teachers reducing the quality of the evidence on which assessments were based in order to cope with the demands on their time. The implications of this experience for teacher education, where unit standards are to be applied, are considered. It is argued that the problems noted in secondary schools will need to be addressed but the costs of compliance are likely to be greater because the teacher education unit standards, on which judgements are required, contain more detailed and numerous performance criteria, many of which will need to be assessed in more than one course, often taught by different departments, and during practical teaching. Three ways in which the difficulties might be resolved are proposed.*

There has been considerable debate in recent years about the New Zealand Qualifications Authority (NZQA) intention to list all New Zealand tertiary qualifications on a single National Qualifications Framework. Every qualification, regardless of its content or vocational focus, would be defined in terms of unit standards to facilitate ease of credit transfer in a much vaunted 'seamless system'. Most of the debate (see for example Irwin, Elley & Hall, 1995; Hager, 1995; and Peddie & Tuck, 1995a) has focused on whether the unit standard system is appropriate for programmes of professional preparation and liberal study where the competence to be tested involves the mastery, integration and application of sophisticated ideas which do not readily lend themselves to being reduced to a series of outcomes in the form required by the Authority. What has not been addressed to date in any detail, is the impact that such an approach is likely to have upon teaching and learning in existing programmes (Peddie & Tuck, 1995b). Nor has any detailed consideration been given to the cost to provider institutions of compliance in terms of required systems, staff time and other material inputs.

Unit standards are now being prepared for teacher education in New Zealand. Since 1993, the Teacher Education Advisory Group (TEAG), an advisory committee of NZQA representing current providers, interested professional organisations and the employers of teachers, has been working its way through a

complex set of negotiations on the matter. To date, a set of domains has been defined and a matrix of unit standard titles exists, together with a set of draft unit standards. Some are already registered on the Qualifications Framework.

Steps are also under way to apply unit standards to teaching in New Zealand secondary schools at the level of non-compulsory education (Year 11 and above). Unit standards were developed in geography and mathematics in 1993-94, and it is intended that all schools will be using unit standards by 1998. In preparation for this, late in 1994, sixty-six secondary schools throughout the country began a trial application of unit standards to the teaching of geography and mathematics to investigate the moderation processes necessary to ensure consistency both within and between provider institutions. The schools involved had to come to grips with a number of broader issues related to using unit standards as the basis for assessing student learning.

In 1995, as a part of the wider moderation trial, one of the writers (Keown) was commissioned by the NZQA to work with six schools in what became known as the Sufficiency Project, to investigate the amount and type of evidence necessary to substantiate that students had met unit standards, and to identify some cost-efficient ways by which this might be achieved. The experience provided an appreciation of some of the costs of compliance when unit standards were applied to existing programmes.

The purposes of this paper are to identify some significant aspects of the experience of the secondary school teachers who applied unit standards to the teaching of geography and mathematics as part of the Sufficiency Project in 1995 and to consider the likely implications for teacher educators who will be required to use national framework unit standards in the future. In the process, attention will be drawn to some likely costs of compliance.

UNIT STANDARDS

A unit standard is defined by NZQA as a statement in outcome terms of critical factors and criteria for the assessment and recognition of achievement. It is said to provide a standard against which students may be assessed and identifies what a student is expected to know, do or understand to demonstrate mastery of a Qualifications Framework unit. All unit standards, regardless of subject or level or complexity, conform to a common format including a unique title, level and credit value, sub-field identification, statement of purpose, entry information, accreditation options, moderation options, special notes, a set of elements (outcomes) and performance criteria.

To use a unit standard to assess learning outcomes, the teacher must devise a suitable learning activity, or use one already prepared by the provider institution, or adapt one of the exemplars supplied by the NZQA. Marking schedules are used to decide whether student responses are up to the required standards and therefore worthy of credit.

From the teacher's point of view, the elements and performance criteria are of particular interest because they have significance for assessment activities. Elements are defined as "the competencies/achievements that must be demonstrated for successful completion of a unit". (NZQA, 1993, p. 25) Each element is accompanied by a series of performance criteria that are defined as

“statements against which the attainment of elements/outcomes is measured” (NZQA, 1993 p. 26). Some performance criteria are also accompanied by range statements. A range statement specifies a range of situations or contexts in which a particular performance criterion must be observed to demonstrate mastery.

It should be recognised that unit standards are not necessarily the names of the courses that make up a programme. It is possible for a unit standard title to coincide with a particular course, but that is not required and tends to be unusual. Thus, for a programme such as pre-service teacher education for primary teaching, leading to the award of the National Diploma of Teaching, there will be a required set of unit standards. How any provider institution structures its programme is up to it, but the provider must be able to demonstrate that every learner who is awarded the qualification has reached the required standard of performance for each unit standard. Consequently, appraisal of performance on some unit standards may occur in more than one topic, unit or course in the overall programme and several unit standards may be achieved in a single topic, unit or course.

This is important because, in theory, educational providers are free to structure their own programmes and teach as they see fit, as long as they can demonstrate achievement of the required unit standards. However, there are also implications for the organisation of assessment within courses and programmes, especially the requirement to have systematic procedures to record the attainment of individual students where assessment of one unit standard takes place in more than one course, perhaps involving different departments, and possibly in different years of a programme.

It is equally important to recognise that unit standards apply to the performance of individual students. Those teaching to any unit standard must record when each student has met every performance requirement for every element, and if range statements are included, every performance criterion must have been demonstrated separately in each situation specified. In theory, a pass may be awarded only when the assessor is confident that all the performance standards have been achieved. There is no provision for a pass to be awarded when the student has met most of the criteria, but not all of them.

Regardless of the specificity of any unit standard, the teacher must still make important judgements, including the quality of performance to be accepted as evidence of mastery, the number of competent performances required, the range of situations expected and the time span during which they must be observed in order to be judged competent. In short, the assessor determines what will constitute a valid package of evidence that competence has been achieved.

THE SECONDARY SCHOOL TRIAL

The six schools in the Sufficiency Project were chosen from the 66 schools in the total trial and were located in the Auckland, Waikato-Bay of Plenty and Canterbury districts and included co-educational and single sex schools, large city schools and provincial and rural schools. Two schools were involved in both geography and mathematics trials and four in trials of either geography or mathematics.

The study involved 16 teachers and 16 classes in 8 departments across 6 schools. In each department, a relatively senior, experienced teacher and a more junior, less-experienced teacher were nominated, one with a class at unit standard level one and the other at level two. Most of the schools nominated teachers who were relatively comfortable with the unit standards process. Each school was visited three times by the research co-ordinator. Every teacher completed a questionnaire and was interviewed on two separate occasions. Five students from each class completed questionnaires and took part in two focus group interviews. Other data were collected in the form of sample assessment activities and recording sheets, and observations of 64 lessons were analysed. The details of the Sufficiency Project have been reported elsewhere (Keown, 1996).

A number of features of the experience of these secondary schools in the trial implementation of geography and mathematics illustrate some of the costs of compliance to the unit standards regime and appear to be of significance for the implementation of unit standards in teacher education.

Teacher Influence on the Final Form of the Unit Standards

The teachers in the trial implementation directly influenced the final form of the unit standards. Although written in 1993-94, the unit standards were still in a draft state when the national moderation trial began early in 1995. When it became apparent that teachers in all 66 trial schools were uncomfortable with some features of the draft documents, their comments were considered before the final forms were approved in July 1995.

The changes made the standards simpler, less wordy, more direct and easier to understand than the originals. Statements of purpose and special notes were simplified by referring to relevant syllabus statements rather than detailing topics and content already set out in syllabuses. Elements and performance criteria were clarified by stating them more directly and performance criteria were expressed in simpler, more readily understood forms. Range statements and merit statements were all eliminated which had the important effect of reducing the number of situations in which mastery of every performance criterion had to be demonstrated. This reduced the assessment demands significantly and put the onus on the teachers to ensure that mastery was demonstrated in a sufficient range of situations, which would confirm for them that the intended standards had been achieved.

The revised structure and writing implied to teachers that they could be trusted to make important professional judgements for themselves. In hindsight, the units are less legalistic than the originals and more user-friendly. The effects on the teachers involved in re-drafting the unit standards documents was a clearer appreciation of what was intended and some sense of ownership of the standards. Many considered the rewritten standards statements easier to apply than the originals, a situation which fostered positive attitudes towards the unit standards and their implementation.

Initial Effects on Teaching

Returning to the six schools under study, there was a tendency for the demands of assessment to greatly influence the way many of the teachers taught, particularly in the early stages of the trial. From the outset, the NZQA intended that the implementation of the unit standards should not require the teachers involved to change their teaching programmes. Their geography and mathematics courses could be taught more-or-less as they had been in 1995, and the new unit standards used to check on learning outcomes. In other words, the unit standards would be applied independently of the curriculum.

In spite of this intention, the demands of assessment soon came to dominate the teaching of many of the project teachers. This was probably influenced by the novelty of a new approach to assessment and the dual assessment system in operation, under which students were also being prepared and assessed for the national qualifications of School Certificate or Sixth Form Certificate. Nevertheless, it became increasingly clear from both observations and the comments of teachers at mid-year (Keown, 1996 p. 175) that more time was being devoted to assessment activities and as a consequence, some reported a reduction in the range of learning activities and corresponding reduction in inquiry learning by students.

It seems that the requirement for the teachers to justify their judgements about student mastery caused them to believe early in the process that the best evidence would be written, and gathered formally. As a result, short assessment tasks, tests and some project work tended to be the order of the day and many teachers noted the effect upon students. However, it is probably more accurate to regard the problem as one of over-testing, rather than over-assessment because later in the year, evidence was gathered in a greater variety of ways and particular students targeted for further evidence.

It is interesting to note that the students had different perceptions of the alleged over-assessment. Late in the year, more than half the 82 students who returned questionnaires (54%) felt the balance between assessment, learning information and practising skills was about right. About a quarter (27%) said they were assessed too frequently and 15% considered that not enough assessment had been done. It was the researcher's interpretation that the latter group was probably from classrooms where the unit standards trial had been accorded a very low priority.

The Need for a Different Approach to Assessment

Assessment in the unit standards required many teachers to learn new and rather complex assessment skills which took time. Two different approaches were evident from the outset. One involved holding regular tests to check on mastery. In a sense, the NZQA expectation that the new unit standards might be applied to the existing curriculum appears to anticipate such an approach which was especially common among the mathematics teachers.

The other approach, more common among geography teachers, involved greater use of ongoing, less formal assessment, based on 'naturally occurring evidence', integrating learning and formal assessment and other forms of informal

assessment. Nevertheless, some geography teachers favoured using the unit standards for summative evaluation purposes and some mathematics teachers supplemented their summative evaluations with data from informal observations and naturally occurring evidence. It seems significant that many teachers broadened their individual repertoires of assessment during the project.

Identifying Standards

Although the unit standards elements and performance criteria identified the student tasks, they said little about qualitative aspects of the standard of performance. Those judgements were left to the teachers which is recognised as a common shortcoming in so-called standards-based systems of appraisal both in New Zealand and in other countries (Eisner, 1995; Hawe, 1995; Irwin, Elley & Hall, 1995). Teachers had to create their own assessment criteria and tests in order to implement the unit standards. Thus, it is hardly surprising that most of the teachers felt pressured by the assessment requirements because in this first year, they were at the same time inventing a new system of assessment and implementing it while also continuing to teach normal loads. An assessment handbook is now available which contains a range of examples which should provide a basis for teachers and schools to develop their own assessment tasks and programmes.

It was also clear that one satisfactory performance is hardly evidence of mastery. Ideally, more than one demonstration should be required, at different times and in different contexts as sufficient evidence that a standard had been met. Many teachers solved this problem initially by accepting, (somewhat unquestioningly), an early suggestion by the NZQA, that three demonstrations of mastery should be required on different occasions before a student might be judged competent on any standard. This generated considerable work for the teachers, regardless of whether the assessments were summative, formal tests, naturally occurring evidence or combinations of these, because three separate assessments were required for every student, together with re-assessments where necessary; and each had to be recorded. A later ruling was that teachers gather only as much evidence of mastery as was needed to show competency.

Further, increasing use was made of naturally occurring evidence, gleaned by observing and listening to students as they performed learning tasks, and in some cases by asking questions and interviewing individuals in informal situations, but hardly 'natural' occurrences. Some teachers also accepted homework task performance as evidence for re-assessment purposes (Keown, 1996 p. 69).

One teacher interviewed by Keown (1996 p. 69) expressed a point of view that became increasingly common as the project continued, especially among the geography teachers. He advocated the use of evidence gathered informally as well as formally, and the integration of assessment with classroom learning activities. His views on authenticity, implicit in the acceptance of evidence from tasks completed out of class and shared student work, were not accepted by all the teachers, especially teachers of mathematics who tended to believe that the only way to assess properly was under test conditions. However, by the end of the project, many of them expressed a desire to relax and broaden the range of

evidence accepted 'in the future', and some planned to be more liberal in 1996. As the year progressed both groups made increased use of oral re-testing.

The teacher's comments also indicate that he had moved from regarding the assessment of competence as simply assessing each student on a series of tasks, to considering the package of evidence of competence (including performance on formal tasks plus class activities and homework) to arrive at a judgement about whether or not he was confident about the evidence as a whole. The literature from overseas suggests that this is a desirable direction of development (Raggatt & Hevey, 1994).

Another mathematics teacher interviewed (Keown, 1996 p. 97) developed a highly efficient system for collating a variety of evidence using a filing cabinet in which she maintained a work profile for every student, which was available to them. Each profile contained an ongoing record of the student's achievements in relation to each unit standard, together with evidence such as a test, homework and other problems completed in class, making up a portfolio of achievement evidence relevant to the standard. Using the work profiles, the same teacher developed an efficient and effective way to identify competence gaps and help students to overcome them.

As the year progressed, many teachers began to think about the quality of evidence along similar lines to Simeonsson and Bailey (1991) who noted three levels of certainty about evidence: 'suggestive evidence', indicating the possibility that the student could achieve competence, 'preponderant evidence', where all indications are that the student has probably achieved competence, but there is not yet sufficient evidence, or evidence of the right type for the teacher to be absolutely sure, and 'conclusive evidence', where there is no doubt that a student is, or is not, competent.

Following this line of thinking, many recognised the practicality of noting those students about whom their current evidence was preponderant and then hunted for further data to be conclusive. For this purpose, some teachers used oral probing of student explanations as a basis for re-assessment and others allowed students to re-work portions of formal tasks in order to make their understanding more explicit. These measures became known within the Sufficiency Project as the 'fix up/re-do' approach. The students involved became categorised by the teachers as 'preponderants' and it is interesting to note that students as well as teachers were of the opinion that to focus on them in this way was fair, given the constraints of time in class (Keown, 1996 p. 63).

Communicating Requirements to Students and Providing Feedback

The teachers recognised the importance of providing good feedback to students but many felt that they had not done well in their first year. This matched the views of students. In hindsight, this was probably inevitable under circumstances where the teachers were still learning about the process while implementing it. However, a number also noted the motivational value of keeping students informed about their progress in relation to the required standards (Keown, 1996 p. 68).

Conversely, others pointed out the difficulty that students experienced in meeting standards when they were not clear about teacher expectations. At the

same time, communicating the nature of the standard was shown to be a time-consuming process that was not really new because lesson analysis showed how teacher talk tended to emphasise this, regardless of whether the teaching related to either unit standards or the School Certificate or Sixth Form Certificate requirements for which the students were being prepared.

Many teachers commented on the difficulty of finding time to develop tasks, mark them, give feedback and offer re-assessment opportunities while trying to maintain their other regular teaching functions. Where choices were made about what could not be fitted in, providing feedback often seemed to become the victim.

It was interesting to note some evidence of students apparently using the first attempt at a task primarily to find out what they had to do in order to demonstrate mastery of the standard. Both students and teachers commented on the fact that some individuals did not really try on the first attempt, knowing that further opportunities to meet the standard would follow.

Tracking

Effective application of the unit standards to the school programmes required systems for the teachers to track the progress of individual students. Reference has already been made to the mathematics teacher who maintained a work profile for every student in a filing cabinet. Other teachers used various forms of recording on paper, frequently involving grids which were ticked to record progress. The need for such measures is readily apparent when it is acknowledged that some elements and performance criteria for one standard for a process might apply to more than one topic taught at different times of the year, and that individual students progress at different rates. Indeed, without an effective tracking system, it is difficult to imagine how any teacher could apply the unit standards effectively.

However, there was a considerable time cost in developing and refining a workable system. The teacher who used work profiles, for example, said that before she could develop a workable system she needed a comprehensive understanding of the unit standards involved and a great deal of practice was necessary before she began to make her system work. All the teachers in the trial worked on the problem of tracking and by the end of the year tracking sheets had been developed with which many teachers felt comfortable. However, these often endorsed a one performance approach, at variance with other ideas developed about sufficiency and competence. Consequently, not all the teachers were equally comfortable with them because it seemed that sufficiency of evidence was being sacrificed in order to achieve a workable system.

Lesson Management

As the year progressed, lessons became increasingly complex. One reason was the necessity for teachers to help students to appreciate the standard or standards to be attained in the current topic. Another was the requirement to offer students opportunities for re-assessment when they failed to demonstrate competency on a given standard, and as has already been noted, some teachers also actively sought

further evidence about the competence of students on whom their evidence to date was only preponderant. Thus, in one lesson, it was not uncommon for teachers to be transmitting new content, signalling performance criteria to students, providing feedback as they discussed problems and working with a group of students preparing for a catch-up assessment or perhaps running a re-assessment test for some in an adjacent room. Under these circumstances, it is not hard to understand why many considered resourcing and time management to be issues of critical importance.

One major issue was the need to provide re-assessment opportunities. Various strategies tried included offering re-sit opportunities during class times (meaning that the students concerned missed out on other work), or at out-of-class times, such as lunchtimes (which some teachers considered unfair on the students as well as the staff involved). Others used fix up/re-do activities to cut down on the number of full re-sits required, and in some cases the examinations associated with the dual assessment system in operation were used as re-assessment opportunities. Regardless of the methods used, they created time and management issues for the teachers and added to their burden of duties.

The Cost of Compliance

These observations suggest that the trial application of unit standards to the teaching of levels 1 and 2 geography and mathematics in secondary schools was completed at considerable cost to the schools which had to adapt existing practices and establish new systems. The price was 'paid' in terms of significant additional inputs of staff time and material inputs, only some of which easily translate into dollar terms. The evidence suggests that the innovation was seriously under-resourced, especially in terms of staff time because most of the teachers involved found themselves working extremely long hours in order to keep up with the changes while also sustaining the other regular aspects of their teaching duties.

Some of the demands were probably short-term in nature and more dramatic than they are likely to be in future applications of unit standards in other subject areas because resource materials to support the implementation of the new approach had not been developed in 1995. It is heartening to note, for example, the recent publication of some resource materials such as the recent draft of a guide to assessment that has been widely circulated for consultation and comment (New Zealand Qualifications Authority, 1996) which will save time for teachers moving to unit standards in the future the most obvious being the extent to which teaching had to be changed to meet the unit standards requirements.

Some of the changes appear both desirable and timely, such as the pressures for teachers to take account of a wider range of evidence, to integrate assessment with teaching, to adopt a more individualised approach, to communicate more about standards and provide improved feedback are clearly desirable trends. However, the observer questioned whether teachers would be able to sustain some desirable evolving practices with the limited human and material resources available. Thus, in some cases, teachers initially adapted to the new requirements by adopting assessment practices that maximised the validity of their judgements about competency, albeit at the cost of excessive personal demands on them and their teaching, it appeared by the end of the year, that some had begun to reduce

those demands by modifying their practices in ways that reduced the validity of their judgements. Some evidence of that was observed, for example, among teachers who reduced the number and/or variety of observations of performance required in order to judge that a standard had been achieved. This suggests that already some excessive costs were being reduced at the expense of quality of assessment.

IMPLICATIONS FOR TEACHER EDUCATION

Currently, 23 unit standard titles have been written for initial pre-service education and detailed unit standards have been registered for 21 of them. They vary considerably in the numbers of judgements to be made about specific competencies. The first five are especially demanding.: 1. *Assess to support student learning*, 2. *Plan and prepare for student learning*, 3. *Create and manage a learning environment*, 4. *Implement planned learning episodes and sequences*, 5. *Demonstrate and apply knowledge of curriculum statements associated with specialisms*. For example, the first one comprises 7 elements, performance on which will be judged according to 22 criteria. The number of required performances increases by a further two when account is taken of a range statement requiring students to demonstrate competence on one criterion separately in regard to diagnostic assessment, formative assessment and summative assessment. This means that 24 decisions about competence are required, but the standard also specifies that assessment on each must be made "within a class programme across at least 3 weeks in at least 2 subject areas or 2 essential learning areas . . ." (Special Note 2). Thus, in reality, 48 judgements must be made about the competencies of individual students in specific situations for one unit standard yielding 20 credit points and credit for the unit should not be awarded unless the individual is judged competent on every one of them.

Apart from the first five unit standards some of the others appear more reasonable in their demands, requiring an average of between 15 and 18 assessments per unit, depending upon whether or not account is taken of range statements. Nevertheless, the overall picture of the requirements is that 21 unit standards that are likely to be the main requirements of a three-year pre-service programme. They will involve recorded judgements about individual student using 311 performance criteria, in between 409 and 516 situations, depending upon whether or not account is taken of range statement requirements. In fact, the number will probably be higher because some performance criteria clearly require a number of separate judgements and some special notes also imply a similar need.

This picture contrasts with the somewhat leaner form of the geography and mathematics unit standards applied in the secondary school trial. Typical examples are a level 2 unit standard (*Explain the nature, patterns and growth of an urban settlement*) with 3 elements and 8 performance criteria, and a level 3 unit standard (*Examine a current geographic issue*) with 3 elements and 6 performance criteria. Neither contains any range statements and the special notes add no further demands. Thus, the numbers of judgements required to decide if an individual has met the standard are 8 and 6 respectively.

The findings of the trial application of unit standards in secondary schools highlights a number of issues to be considered by those who anticipate providing teacher education under a unit standards regime. Questions need to be asked about likely compliance costs to providers in the form of required changes to systems, the likely effects upon teaching and learning and additional demands upon staff.

The Form and Number of Unit Standards

The experience in secondary schools highlights some advantages of expressing unit standards in a simple form, with clear intentions and a limited number of elements and performance criteria per element. This avoids increasing the number of assessment decisions required of teachers. There should be extremely good reasons for the inclusion of any range statements because, as has been illustrated, every additional range statement item multiplies the number of assessment decisions and it needs to be recognised that in order to arrive at valid judgements, every performance decision requires evidence of mastery to be gathered and recorded on several different occasions.

The specificity of most of the performance criteria and range statements reinforces a mistaken impression that what are called unit standards are in fact standards of performance. Regardless of how much detail is built into them, the question of what constitutes quality of performance must still be answered by assessors and it is the view of the writers that the removal of such detailed specifics as range statements helps make it clear that those who assess to the standards must still make decisions about the quality and quantity of performance, taking account of the general intentions of the standard statement. The question of what is necessary and sufficient evidence that a standard has been met, especially in the qualitative sense, must be answered by assessors.

The NZQA intends every unit standard to have an initial 'shelf life' of only two years before it is reviewed. In spite of this, the almost accidental occurrence by which the secondary school teachers were able to influence the drafts before they were finalised, would be desirable in the development of all unit standards. None of those involved in writing the teacher education unit standards, or those who advised them (including one of the writers of this paper) has any experience of working in a unit standards regime so it is almost inevitable that there will be shortcomings which those preparing to implement the standards may identify as they plan the details of their implementation. If that is the case, there should be opportunities for changes to be made at an early stage, if shortcomings are revealed, rather than to implement flawed documents. Ideally, as in the secondary schools, general implementation should be preceded by a trial.

Relating Unit Standards to Programmes

The trial application of unit standards in secondary schools was relatively simple compared with what is likely to be required of teacher education programmes. In the schools, unit standards were applied to learning in independent subjects, taught by separate departments. The teachers were also accustomed to working under the constraints of relatively explicit national syllabuses and examinations.

The situation facing the providers of teacher education is dramatically different. Many of the unit standards do not coincide with existing courses, subjects or even the academic territories of teaching departments. They relate to ideas and skills that are developed and applied across a range of courses taught by different departments. The learning activities involved include both campus and school-based activities and it is likely that many will apply to more than one year of any student's programme. This means that the first stage of implementation should involve a careful mapping exercise to identify where in the programme an individual student might be assessed on each performance criterion. Having a clear overall map of all the possible assessment points is the logical starting point.

The second step should be to identify the main evidence gathering point for each performance criterion. This would be the stage of the programme when every student would be first assessed on a given criterion. The potential assessment points that follow would then become chances for further demonstrations of mastery as evidence of competence, or opportunities for re-assessment, when that is required.

Tracking Student Attainment

The secondary teachers were required to track the attainment of individual students in one course under their supervision for the year. This may appear to be complex, but it is the writers' view that the demands were simple compared with the complex tracking that will be required by the teacher education unit standards. For the best of educational reasons, many of them focus on knowledge and skills relevant to several aspects of teaching. For example, a standard requiring a student to be able to 'assess to support student learning' requires students to demonstrate mastery of performance criteria in more than one curriculum area, and it is common for courses on the teaching of different subjects to be taught by staff in different departments. Judgements about mastery of one performance criterion may require assessments in more than one course, and in practicum. In some areas, it may also be necessary to gather evidence over the full period of pre-service studentship. In the light of difficulties experienced by the secondary teachers who had to evolve workable procedures to track and document student learning during one school year, this must be accorded priority by teacher educators who assess to unit standards. Devising and operating an effective system for a whole programme will be no mean undertaking.

It is possible that students themselves may help to manage the tracking by maintaining their own portfolios of evidence. It would keep them informed of their own progress, as is recognised in teacher education programmes that already use the technique. However, the nature of such student involvement will need to be carefully considered because it must be practically feasible, and consider regulatory and legal implications at a time when dissatisfied students are becoming increasingly inclined to resort to legal redress. Ultimately, it is the provider institution that must decide whether or not any unit standard has been met, not the student, and under a unit standards regime, it will be important for the evidence on which judgements were made to be fully documented, especially in cases where standards were not achieved.

Where student learning on one unit standard is assessed in different courses, it will also be important for provider institutions to be quite clear about who will make the final decisions about competency. Where a given unit standard, or even one element of a unit standard, applies to only one course, the customary practice, by which the course teachers also act as the examiners may apply. Normally, they are simply expected to meet the set of quality control measures required of all departments. However, the situation is less clear when the evidence is gathered in courses taught by more than one department. Where this happens in teacher education, it will be important to be quite clear about who is the examiner and what part will be played by the lecturers from the other courses involved, including associate teachers and visiting lecturers, for practicum assessment.

Sufficiency of Evidence

The demand for lecturers to make judgements about whether unit standards have been mastered will probably lead them to work through the same sequence of activities required of the teachers in the Sufficiency Project. First, following careful study of the performance criteria in the unit standards, assessment tasks will need to be devised. Then examples of student work must be gathered and assessors will need to debate which work meets the standards, until they arrive at acceptable performance. In the process, it will be important to focus on qualitative as well as quantitative aspects of performance.

Besides considering when and why a single performance may be judged competent, it will also be important to consider the number of demonstrations necessary, the variety of situations in which evidence of mastery should be expected and when in the programme an individual may be judged to have achieved every standard. As in the secondary schools, these decisions that have a strong bearing upon the validity of the judgements made about competence will need to be balanced against the relative costs of gathering a range of evidence compared with one-off observations. It is the view of the writers that the very large number of judgements required by the present teacher education unit standards runs the risk of encouraging providers to make judgements based on limited evidence at the expense of reduced validity.

Methods of Assessment

In the secondary school trial, as in overseas research, it seemed that the teachers with a range of formal and informal assessment procedures at their disposal were those best equipped for the task of assessing using unit standards. The situation in teacher education is likely to be exactly the same. Being able to use a variety of formal and informal methods to gather evidence of mastery on required tasks will have the twin merits of reducing time spent on formal assessment and making it possible for lecturers to make ongoing assessment a natural part of their teaching.

The current teacher education providers are probably in a better position to do this than the secondary teachers who were studied. Many lecturers in teacher education already use a mixture of formal and informal methods of evaluation and many of their preferred learning activities will also adapt readily to gathering informal evidence of mastery. However, recent trends towards creating larger

classes in order to teach more cost-efficiently, in the face of reducing funding, could make it more difficult to sustain some of those activities in the future and even if they can be sustained, may also make it more difficult for staff to gather assessment data in class. In other words, increasing class sizes could force a greater reliance upon formal methods of assessment.

Moderation

A further challenge to teacher education posed by the unit standards lies in the area of moderation. As in the secondary schools, this will involve moderation to ensure consistency within provider institutions and between them. It has already been noted that there is potential in teacher education for assessments of mastery on one performance criterion to be required in more than one course. If this happens, institutions will need to moderate the assessment procedures and standards of different departments and teaching teams in ways which have not been demanded in the past. The secondary school trial indicated how teachers in one department are able to check on and maintain comparable standards through normal professional contacts. However, some of the teachers said that the moderation of standards between schools appeared more haphazard and required much more effort and resourcing to make it effective. In a number of instances, teachers expressed concerns that they required higher standards of their students than those required by their moderators so that they believed lesser performances might be accepted elsewhere.

It is likely that moderation between providers will pose similar problems for teacher education. Moreover, assessments on one unit standard in courses taught by different departments (or even different teams) of a large provider institution, are likely to create another internal moderation issue. As in the case of inter-institution moderation, the issue is not new. However, addressing it is likely to cost precious resources because it will involve both time and staff development, especially in the initial years of implementation.

Re-assessment

One important feature of the unit standard system is its provision for students who fail to achieve a unit standard to be re-assessed at a later stage of the course. In the secondary schools, this had important implications for the work of the teachers because it complicated their teaching later in the year. It also meant that the demands of gathering further evidence about the competency of students seeking re-assessment, for which there was only limited time available, required them to decide which students should be accorded priority. Thus, they elected to focus upon students who were considered close to achieving mastery, the preponderants.

Re-assessment is not a new practice in teacher education. Most current programmes offer some opportunities, although usually as concessions to students. Under a unit standards regime, however, re-assessment becomes a right. It was noted in the secondary school trial, that some students did not try very hard to pass at the first assessment opportunity because they knew a further attempt would be forthcoming. This phenomenon has also been observed among

American students in an outcome-based system of teacher education with some features similar to that anticipated in New Zealand under unit standards (Towers, 1994 p. 627). Towers also noted that while the Minnesota scheme appeared to support low achievers, the demands on the teaching staff under a regime of reducing funding meant that those students tended to benefit at the expense of higher achieving students.

It is anticipated that as in secondary schools, managing the re-assessment demands will become a key issue that calls into question both the right of students to continue to be re-assessed on a standard which they failed to achieve on a number of previous occasions, and the extent to which an individual should be permitted to proceed in a programme while pre-requisite competencies have not been achieved. Again, these are not new issues in teacher education and each provider institution will need to deal with them in its own way. It is suggested that the key principle to be remembered when establishing policies and practices is that the unit standards are being applied to the programmes of provider institutions and that while some flexibility may be possible under the unit standards regime, it should be within the powers of provider institutions to establish additional conditions of studentship which apply to the achievement of the qualification in question. In practice, it should remain appropriate for a provider to require the passing of certain courses (or the achievement of certain standards) as a pre-requisite to enrolling for others, and to limit the extent to which an individual may proceed in a programme without passing earlier courses. Programme completion may be required within a limited period.

If students are forced to leave a programme because of unsatisfactory progress, it will remain possible for them to transfer their unit standard credits to another programme, but only if the new provider is willing to enrol them.

Resourcing

One of the most serious implications for tertiary providers of teacher education working in a unit standards regime will lie in the area of resourcing. This was evident in the secondary school trial where implementing the new system imposed excessive workloads. The problem was caused partly by the demands of implementing a new system under the conditions of dual assessment referred to earlier. While it may be possible for teachers to streamline their procedures in the future, assessment under a unit standard regime such as that proposed by the NZQA, is always likely to be more demanding of time and skill than the previous system. To make the unit standards work well in secondary schools will require more generous staffing, staff development and improved information handling systems, all of which are dependent upon funding. Similarly, the establishment of adequate moderating procedures to ensure consistent standards between schools will also require significant human resources.

Thus, it is the view of the writers that the initial implementation of unit standards in teacher education will make significant additional demands on staff for the reasons noted in the secondary schools. It will be a time-consuming exercise, more so if standards are to be applied across courses and years of the same programme. The initial establishment of workable and credible tracking procedures will be especially demanding of both staff time and information

systems but many of the other demands for resources are likely to be ongoing. These are important considerations, because the likely new demands will be made at a time of progressively decreased funding. If appropriate resources are not provided, the experience in the secondary schools suggests that there is a danger that lecturers may trim their systems of assessment to match the resources available. The most obvious way of doing that will be to rely increasingly on one-off assessments as the basis for judgements about mastery, rather than the more desirable packages of evidence, thereby missing one of the key intentions of using unit standards as a basis for assessment.

CONCLUSION

Some features of the unit standards regime as observed in the secondary schools in the Sufficiency Project have the potential to enhance the quality of any educational programme. They include the desire to base assessment on a wide range of relevant evidence, to integrate assessment more fully with teaching, to adopt an increasingly individualised approach to education, to give students a clear idea of what they are expected to achieve and to provide them with improved feedback on their progress. However, under the unit standards regime, the achievement of those potential benefits comes with a set of compliance costs which are not always fully appreciated. The costs include the requirements for teachers to master new assessment techniques, develop ways of tracking and recording the achievements of students, manage the demands of re-assessment and integrate these functions with their other teaching duties. Thus, the successful implementation of a unit standards regime has significant resource implications, which appeared to be inadequately addressed during the period of the Sufficiency Project. If these matters are not addressed in the future, teachers may reduce the costs of compliance by basing their judgements about competence on reduced evidence, thereby also reducing the educational validity of their assessments.

When NZQA unit standards are applied to pre-service teacher education, the same issues will need to be addressed. However, the more detailed requirements of the teacher education unit standards, together with the fact that many of them will involve assessment in more than one course, taught by staff in different departments, over more than one year of the programme, including assessments during periods of practicum, appears likely to increase the compliance costs significantly. As in the secondary schools, one way of reducing those costs will be to reduce the range of evidence gathered as the basis of assessment, with the consequence of reducing the validity of the judgements. The temptation might be to reduce the quality of the assessment, simply to fulfil the unit standards requirements.

These trends suggest that there are three possible ways in which institutions might respond to the application of unit standards to teacher education. One is to seek the additional resources that will clearly be required for the standards to be implemented on a valid and sustainable basis. The second would be to simplify the standards themselves to reduce the cost of compliance. The third would be to seek the educational benefits of improved systems of assessment without the

restrictions and the costs of compliance of the unit standards regime by registering provider qualifications on the National Curriculum Framework.

REFERENCES

- Eisner, E.W. (1995). Standards for American Schools. *Phi Delta Kappan*, 76(10) 758-764.
- Hager, P. (1994). Is there a cogent philosophical argument against competency standards? *Australian Journal of Education*, 38(1), 3-18.
- Hawe, E. (1995). Standards based assessment and the New Zealand Draft Technology Curriculum. In Peddie, R. and Tuck, B. (Eds.). *Setting the Standards*, pp. 120-139. Palmerston North: Dunmore Press.
- Irwin, M., Elley, W. and Hall, C. (1995). *Unit Standards in the National Qualifications Framework*. Auckland: The Education Forum.
- Keown, P. A. (1996). *Walking the Sufficiency Ridge*. The final report of the NZQA Sufficiency Issue Research Project. Hamilton: Department of Social, Physical and Health Education, University of Waikato.
- New Zealand Qualifications Authority (1993). *Setting and Maintaining Standards for the National Qualifications Framework: A Guide for National Standards Bodies*. Wellington: NZQA.
- New Zealand Qualifications Authority (1996). *Learning and Assessment: A Guide to Assessment for the National Qualifications Framework*. Wellington: NZQA.
- Peddie, R. and Tuck, B. (Eds.). (1995a) *Setting the Standards*. Palmerston North: Dunmore Press.
- Peddie, R. and Tuck, B. (1995b). Fruitful Change: Concluding Comment. In Peddie, R. and Tuck, B. (Eds.). *Setting the Standards*. (pp. 209-204) Palmerston North: Dunmore Press.
- Raggatt, P. and Hevey, D. (1994). Do I have enough evidence? *Competence and Assessment* 27,10-13.
- Simeonsson, R.J. and Bailey, D.B. (1991). Evaluating programme impact: Levels of certainty. In D. Mitchell, and R. Brown, (Eds.). *Early Intervention Strategies for Young Children With Special Needs* (pp. 280-296). London: Chapman-Hall.
- Towers, J.M. (1994). The perils of outcome-based teacher education. *Phi Delta Kappan*, 75(8), 624-627.