Assessing student teachers’ performance in practicum

Mavis Haigh & Bryan Tuck

Auckland College of Education

Auckland, New Zealand

After completing a degree in relevant content areas prospective high school teachers in New Zealand typically enrol in a one-year Diploma at a College of Education. The year is intensive and a critical component is supervised teaching within selected high schools. The student teacher is attached to an experienced teacher, an Associate Teacher, who accepts responsibility for the day to day supervision of the student teacher's teaching experience. During each attachment a Visiting Lecturer for the College of Education also observes the student teaching. This study examines the levels of agreement between the Associate Teacher and Visiting Lecturer’s evaluations of 150 student teachers teaching experience on their third and penultimate school placement. The analyses examine correlations between the judge’s holistic ratings of teaching competence and ratings of specific teaching competencies. The values of the correlations vary dramatically and significant differences exist between the mean ratings of the Associate Teachers and the Visiting Lecturers. These differences are discussed. Structural analyses of the rating instruments are reported and the implications of the results for the assessment of practicum in pre-service teacher education programmes identified.

Teaching practicum is a central element in most pre-service teacher education programmes. However the assessment of students' competence during practicum remains problematic. A number of issues remain contentious, e.g. the tension between the different purposes of assessment (Fish, 1995; Thompson, 1999), the impact of context on practice (Maloney, 1998); the tension between message and saving face (Wajnryb, 1996); who is to be the arbiter and definers of good practice (Fish, 1995); and competence versus competencies (Fish, 1995; Gibbs and Aitken, 1996). Additionally, debates about the assessment of the practice of student teachers often reflect ongoing philosophical debates about the nature of teacher education (Brown, 1996) and traditional barriers between teachers and academics (Groundwater-Smith, 1997).

The development of sets or lists of teacher competencies, dimensions or standards is a world-wide phenomenon, embedded in a much larger 'master narrative' of economic, social and political issues, and often a response to government and community concerns over falling standards and crises in education (Ballantyne, Thompson, and Taylor, 1998). In the United States there are the "standards" of the Interstate New Teachers Assessment and Support Consortium (Ambach, 1996), the extensive work of the National Board for Professional Teaching Standards (see Jaeger, 1997, 1998) for an evaluation of the assessment procedures), and the Praxis series of assessments developed by Educational Testing Service (Dwyer and Villegas, 1992), to name only a few of the many. The United Kingdom has frameworks of competencies for the accreditation of teacher training courses (Barton and Elliot, 1996) and in Australia there is the National Competency Framework for Beginning Teachers, which came out of the three year research and development project of the National Project on the Quality of Teaching and Learning. In New Zealand there have
been a series of semi-independent lists identifying the characteristics of the capable or satisfactory teacher: the Teacher Registration Board (1997) has published criteria for identifying the "satisfactory teacher", the New Zealand Qualification Authority has published "standards" of teaching (Gibbs and Munro, 1993; McGrath, 1996), the Education Review Office (1998) has published the defining characteristics of the "capable teacher", and last but not least the Ministry of Education (1998 and 1999) has recently published "standards" for teachers and administrators in primary and secondary schools. The published models of good teaching vary dramatically in terms of the number of dimensions or domains they identify.

The plethora of publications if nothing else provides evidence that the defining characteristics of the good teacher are both problematic and contested. As Stones (1994, p 236) commented, "consensus on the critical attributes of competent teaching does not exist". At one level no one is going to disagree with Shulman (1986) that a teacher needs to know about the subject matter, to know a variety of general instructional strategies, and to know about the specific strategies necessary for teaching particular subject matter. Most of us would also agree that the good teacher transforms curriculum goals and guidelines in such a way that a particular student is able to master and understand the related content. However, the movement from this level of rhetoric to assessment procedures enabling us to identify good teaching in practice is a process characterised by fuzzy logic.

The first author inherited a set of written criteria used to assess the competence of pre-service secondary teacher education students during practicum. The dimensions of the assessment protocol did not appear to be based on any articulated theory of good teaching practice, and there were significant doubts about the extent to which the various groups of stakeholders had a shared understanding of the standards implied in the criteria. In education what we label as standards are socially constructed and frequently fuzzy (Sadler, 1987) and require the shared understanding of a construct in a community of practice (Wiliam, 1996). There was good reason to believe that the student teachers, the lecturers responsible for supervising teaching experience and the supervising teachers in the schools did not constitute a community of practice in Wiliam's sense.

The study reported in this paper looked at the degree to which Associate Teachers (supervising teachers in the schools) and Visiting Lecturers (tertiary institution lecturers) have a shared understanding of what constitutes good teaching as they assessed student teacher competence during teaching practicum.

**METHOD**

**The Participants**

**The student teachers:** The student teachers whose practicum was being assessed were enrolled in a diploma of teaching for secondary teachers. Entry to the diploma is restricted to graduates. Many had spent considerable time in work between graduation and their entry into teacher education. Their average age was 30. During the 1997 year each student teacher had four block practicum placements. Their practicum experiences were closely linked with their curriculum studies and they were required to pass both the college and school based components of the curriculum studies courses. The college-based lecturers who visited them for observation and assessment in the schools were, in the main, their curriculum studies lecturers.

**The Associate Teachers:** The classroom teachers supervising the students for their school-based practicum are called Associate Teachers. Associate Teachers worked closely with the student teachers helping them to develop as teachers and, at the end of the block
placement, completed an assessment report form (see Appendix 1). The Associate Teachers were required to discuss their decisions and comments with their student teachers and both had to sign off the report. The student teacher’s signature did not necessarily indicate agreement with the Associate Teacher’s comments, merely that they have seen and discussed the report with the Associate Teacher.

**The Visiting Lecturers:** Lecturers at the College of Education visit the students teachers to observe their teaching within their Associate Teacher’s class. The typical lessons ranged in time from 40 to 60 minutes. Each visit included aspects of clinical supervision (Acheson and Gall, 1980), with a pre-observation discussion of the lesson, observation and with post-teaching reflection and goal setting occurring during a dyadic or triadic discussion about the observed lesson. Follow-up visits were negotiated where necessary. Visiting Lecturers were expected to make notes of aspects of the observed lesson and the ensuing discussion. Normally they did not complete an assessment report formatted in a similar way to that used by the Associate Teacher.

**The Assessment Instrument**

The assessment instrument used in this study was the regular Associate Teacher report form (see Appendix 1). On this form there was opportunity for the assessor to indicate the context of the assessment, providing information about the class level, subject, and an indication of the assessor’s perception of the degree of difficulty of management of the class. Assessors were then asked to comment on the student teacher’s professionalism. The main section of the report focused on the student teacher’s competence under five dimensions. The purpose of this section was to identify for the student and college staff those areas that should be a focus for future development. The five general dimensions were:

- Contextual knowledge for teaching
- Planning and organising for teaching
- Managing learning – communicating to enhance learning
- Managing learning – promoting positive outcomes
- Assessing learning

Each of these dimensions had between three and nine indicators. The teachers and lecturers rated the competence of the student on each indicator using the following levels:

" Clear strength, exceeds expectations"

" Acceptable standard"

" Recommended goal for development on next teaching experience"

" Essential goal for next teaching experience"

The raters also indicated the student teacher's overall level of achievement on the practicum. There were five levels for the assessor to choose from ranging from an indication that the student had "serious limitations, suitability for teaching questioned" to "the student was very capable and consistently exceeded expectations". Finally the assessor’s perceptions of the degree of difficulty that a student would have in managing that particular class was rated on a four point scale: " very difficult", " difficult", "reasonable", " straight forward".
Procedures

The Associate Teachers completed the assessment of the student teacher at the end of a practicum of four weeks. They received no special instructions for either the completion or return of the form as the assessment was part of their normal role as associate teachers. The Associate Teachers reported that on average their ratings were based on 10 hours of observation of the student teaching.

A number of meetings were undertaken with the college lecturers to ensure that they understood the reasons for the research and the nature of their role. After their next school visit to observe the student teaching they completed an assessment of the student teacher using the same form as the Associate Teacher.

RESULTS

Correlations between judgements of Visiting Lecturers and Associate Teachers

Table 1 sets out the correlations between the judgements of the Lecturers and the Associate Teachers. There are a number of correlations that are not significantly different from zero, e.g.

"Understands issues related to cultural diversity and equity areas"

"Stresses key or difficult points"

"Instruction giving is clear and direct"

"Ensures physical safety and health of students"

"Gives effective individual guidance"

and the three skills making up the category "Assessing Learning".

The highest levels of agreement, and even here with one exception they are not particularly high, occur on the following:

"Speaks clearly, audibly and fluently" (r=.62)

"Projects enthusiasm for subject" (r=.41)

"Uses language which is understood by students" (r=.38)

"Understands the role of language in effective teaching and learning areas" (r=.39)

"Considers prior and future learning" (r=.35)

"Uses variety of learning activities and teaching strategies" (r=.37)
"Manages class with confidence and purpose" (r=.36)

Despite the relatively low levels of agreement between the ratings of the Visiting Lecturer and Associate Teacher on the level of competency displayed by the student on the 30 specific indicators there was a modest level of agreement (r=.54) on their overall rating of the student’s Teaching Experience.

**Do Visiting Lecturers and Associate Teachers have similar standards?**

Table 1 provides the mean ratings of the students by the Visiting Lecturers and the Associate Teachers on each of the 30 indicators. The ratings of the Visiting Lecturer are with few exceptions "harder" than those of the Associate Teacher. The average of the ratings of the Lecturers (3.9) for "Overall Teaching Experience" is also significantly lower than the average rating of the Associate teachers (4.2). Table 2 sets out a cross-classification of Visiting Lecturer's and Associate Teacher's ratings of the students "Overall Teaching Experience". The distribution is skewed, e.g. seven students received a rating from the Lecturer that was two points lower than the one they received from their associate, 55 students received a rating from the lecturer one point lower, 74 received identical ratings and 14 received a rating from the lecturer one point higher than that of their associate.

**A structural test of the model of student teacher competence**

A principal factor analysis with orthogonal rotation (varimax) was undertaken of the ratings of students by the Associate Teachers. The Associate Teachers' ratings were used because many of the Visiting Lecturers did not complete some of the ratings on Assessing Learning. The variables consisted of the 30 specific competencies of the 1997 Teaching Experience report form. The rotation was limited to the first five factors extracted in an attempt to model the five categories in the original model of teaching experience: 1 Contextual Knowledge for Teaching, 2 Planning and Organising for Teaching, 3 Managing Learning-Communication to Enhance Learning, 4 Managing Learning-Promoting Positive Outcomes, and 5 Assessing Learning.

The initial extraction identified seven factors with eigen values above 1.0. A scree plot suggested a substantial change in the gradient between the fifth and sixth factor. The fifth factor had an eigen value of 1.40, whereas the eigen value of the sixth factor was 1.18, and the seventh 1.07. In the following discussion, loadings on the orthogonal factors are in parentheses, with the discussion limited to items with loadings equal to or greater than 0.50.

There was reasonable strong confirmation of the category Contextual knowledge for teaching.

"Has relevant/appropriate knowledge of subject areas" (0.66)

"Has appropriate understanding of the NZ curriculum" (0.74)

"Understands issues related to cultural diversity and equity areas" (0.53)
However the other skill in this area- "Understands the role of language in effective teaching ..." did not load on the same factor as the above three items.

There was not strong support for the grouping of the skill statements in the second area- Planning and Organising for Teaching. Three of the seven skills in this category- "Appropriate balance between teacher and student activity" (0.52), "Sequential order, timing and pace clearly planned" (0.59), and "Uses variety of learning activities and teaching strategies" (0.63)- loaded on the same factor as the following skills, that were not originally part of this category:

"Stresses key or difficult points" (0.57)

"Instruction given is clear and direct" (0.51)

"Manages class with confidence and purpose" (0.68)

"Encourages student progress through feedback..." (0.69)

"Manages behaviour in ways that promote positive self perceptions" (0.66)

It would seem that the Associate Teachers do not discriminate between prior planning as such and the organisation and management of the teaching session itself.

There was mixed support for Managing Learning-Communicating to Enhance Learning. Four of the 9 skill statements loaded on a common component, and these were:

"Speaks clearly and audibly" (0.71)

"Uses language which is understood by students" (0.83)

"Clearly communicates lesson direction to students" (0.55)

"Instruction given is clear and direct" (0.57)

and another skill initially categorised under Contextual knowledge for Teaching also loaded on this component, i.e.

"Understands the role of language in effective teaching..." (0.50).

Clearly this factor has to do with communication.

There was also mixed support for the category Managing Learning-Promoting Positive Outcomes. Four out of the seven original statements loaded on a single factor, and these were:

"Creates/uses effectively a positive learning environment" (0.51)

"Ensures physical safety and health of students" (0.50)

"Manages behaviour in ways that promote positive student self-perception" (0.66)

"Manages class with confidence and purpose" (0.51)
There was clear support for the category Assessing Learning with all three skill statements loading on the one factor:

"Chooses appropriate assessment methods" (0.84)

"Relates assessment to planning and teaching" (0.77)

"Uses appropriate methods for reporting to students" (0.75)

plus one other statement from outside this category

"Lesson plans state learning outcomes clearly and concisely" (0.59)

CONCLUSIONS AND IMPLICATIONS

The levels of agreement observed between the judgements of Visiting Lecturers and Associate Teachers are somewhat less than one would hope for. Not only are the correlations between the ratings of the judges low, but on average the Visiting Lecturer is a harder judge than the Associate Teacher.

There are a number of explanations for these differences. The most obvious is that the Visiting Lectures and Associate Teachers do not have a shared understanding of the meanings in practice of the indicators, i.e. they are not members of the same "community of practice" (Wiliam, 1996) or "guild" (Sadler, 1989). However they are also not observing the student within the same context. The Visiting Lecturer generally observes the student once over a period of one hour whereas the Associate Teacher's judgements are the culmination of observations they have of the student teaching over a four-week period. It is difficult to explain the "Harder" standards of the Lecturer as a function of differences in sampling, as it is reasonable to assume that the student teacher will be putting on a prepared peak performance for the visit of the lecturer. If anything the judgements of the Associate Teachers should have been harder as they have observed the everyday activity of the student.

The low level of concurrence between Associate Teacher and Visiting Lecturer rankings on many of the 30 competencies and the differences in overall grading scores indicates a need for professional development designed to enhance our shared understandings of what constitutes good practice. In what ways can ACE lecturers and Associate Teachers work together to enhance this shared understanding without adding considerably to the work load of Associate Teachers, school-based college student teacher co-ordinators or lecturers?

The clustering of items on the factorial analyses also raises questions about our model of good practice. Changes to the assessment process need to take into account the analyses in the current study, and the dimensions used by such agencies in New Zealand as the Teachers Registration Board, ERO, PPTA, and NBPTS and NCATE in the USA and their equivalents in the UK and Australia. It is not sufficient for us to reflect on the dimensions of good practice. We must also reflect on the approaches we take to assessment. At one extreme there is the technicist approach exemplified by our rating scales and, at the other there is the use of generic scoring rubrics. Whatever the approach it is essential that students, lecturers and teachers have a shared understanding of what is involved.
The issues and questions generated have been complicated by a new modularised qualification – the Graduate Diploma of Teaching (Secondary). Teaching experience has now been restructured around two stand alone Practicum modules each with a small number Learning Outcomes and associated Performance criteria. We are never going to get this right, but what is essential is that all of those involved continue to reflect on the assessment process and its consequences for the learner teacher.

REFERENCES


Wajnryb, R. (1996). The pragmatics of feedback; Supervision as a clash of goals between message and face. In A. Yarrow, J. Millwater, S. De Vries, & D. Creedy (Eds.), *PEPE Research Monograph: No.1*. Brisbane: Queensland University of Technology

Table 1. Mean ratings of Visiting Lecturers and Associate Teachers on 30 Aspects of Teaching Competency, Ability to Manage Class Level, and Overall Assessment of Teaching Experience (Note that the numbered competencies refer to the 1997 Associate Teacher report form - see Appendix 1)

<table>
<thead>
<tr>
<th>Variable</th>
<th>Lecturer M</th>
<th>Associate M</th>
<th>correlation</th>
<th>N</th>
</tr>
</thead>
<tbody>
<tr>
<td>Context 1</td>
<td>3.20</td>
<td>3.22</td>
<td>.34**</td>
<td>&gt;150</td>
</tr>
<tr>
<td>Context 2</td>
<td>3.21</td>
<td>3.18</td>
<td>.29**</td>
<td>&gt;150</td>
</tr>
<tr>
<td>Context 3</td>
<td>3.38</td>
<td>3.38</td>
<td>.10</td>
<td>81</td>
</tr>
<tr>
<td>Context 4</td>
<td>3.17</td>
<td>3.25</td>
<td>.39**</td>
<td>&gt;150</td>
</tr>
<tr>
<td>Planning 5</td>
<td>3.19</td>
<td>3.34</td>
<td>.35**</td>
<td>&gt;150</td>
</tr>
<tr>
<td>Planning 6</td>
<td>3.15</td>
<td>3.33</td>
<td>.35**</td>
<td>&gt;150</td>
</tr>
<tr>
<td>Planning 7</td>
<td>3.18</td>
<td>3.36</td>
<td>.22**</td>
<td>&gt;150</td>
</tr>
<tr>
<td>Planning 8</td>
<td>3.12</td>
<td>3.18</td>
<td>.18*</td>
<td>&gt;150</td>
</tr>
<tr>
<td>Planning 9</td>
<td>3.17</td>
<td>3.47</td>
<td>.37*</td>
<td>&gt;150</td>
</tr>
<tr>
<td>Planning 10</td>
<td>3.40</td>
<td>3.55</td>
<td>.19*</td>
<td>&gt;150</td>
</tr>
<tr>
<td>Planning 11</td>
<td>3.27</td>
<td>3.38</td>
<td>.26**</td>
<td>&gt;150</td>
</tr>
<tr>
<td>Man.-Com.. 12</td>
<td>3.45</td>
<td>3.33</td>
<td>.34**</td>
<td>&gt;150</td>
</tr>
<tr>
<td>Man.-Com.. 13</td>
<td>3.16</td>
<td>3.18</td>
<td>.62**</td>
<td>&gt;150</td>
</tr>
<tr>
<td>Man.-Com.. 14</td>
<td>3.37</td>
<td>3.34</td>
<td>.38**</td>
<td>&gt;150</td>
</tr>
<tr>
<td>Man.-Com.. 15</td>
<td>3.06</td>
<td>3.17</td>
<td>.18*</td>
<td>&gt;150</td>
</tr>
<tr>
<td>Man.-Com.. 16</td>
<td>3.04</td>
<td>3.16</td>
<td>.12</td>
<td>148</td>
</tr>
<tr>
<td>Man.-Com.. 17</td>
<td>3.13</td>
<td>3.13</td>
<td>.13</td>
<td>&gt;150</td>
</tr>
<tr>
<td>Man.-Com.. 18</td>
<td>3.19</td>
<td>3.32</td>
<td>.41**</td>
<td>&gt;150</td>
</tr>
<tr>
<td>Man.-Com.. 19</td>
<td>2.96</td>
<td>3.12</td>
<td>.25**</td>
<td>&gt;150</td>
</tr>
<tr>
<td></td>
<td>Rating 1</td>
<td>Rating 2</td>
<td>Rating 3</td>
<td>Rating 4</td>
</tr>
<tr>
<td>----------------</td>
<td>----------</td>
<td>----------</td>
<td>----------</td>
<td>----------</td>
</tr>
<tr>
<td>Man.-Com.. 20</td>
<td>3.14</td>
<td>3.38</td>
<td>.29**</td>
<td>&gt;150</td>
</tr>
<tr>
<td>Man.-Posit. 21</td>
<td>3.22</td>
<td>3.40</td>
<td>.28**</td>
<td>&gt;150</td>
</tr>
<tr>
<td>Man.-Posit. 22</td>
<td>3.34</td>
<td>3.48</td>
<td>.02</td>
<td>60</td>
</tr>
<tr>
<td>Man.-Posit. 23</td>
<td>3.21</td>
<td>3.29</td>
<td>.36**</td>
<td>&gt;150</td>
</tr>
<tr>
<td>Man.-Posit. 24</td>
<td>3.12</td>
<td>3.39</td>
<td>.20*</td>
<td>&gt;150</td>
</tr>
<tr>
<td>Man.-Posit. 25</td>
<td>3.22</td>
<td>3.47</td>
<td>.07</td>
<td>&gt;150</td>
</tr>
<tr>
<td>Man.-Posit. 26</td>
<td>3.15</td>
<td>3.26</td>
<td>.30**</td>
<td>&gt;150</td>
</tr>
<tr>
<td>Man.-Posit. 27</td>
<td>3.33</td>
<td>3.55</td>
<td>.24**</td>
<td>&gt;150</td>
</tr>
<tr>
<td>Assessment 28</td>
<td>3.31</td>
<td>3.43</td>
<td>-.16</td>
<td>38</td>
</tr>
<tr>
<td>Assessment 29</td>
<td>3.40</td>
<td>3.39</td>
<td>-.11</td>
<td>30</td>
</tr>
<tr>
<td>Assessment 30</td>
<td>3.65</td>
<td>3.39</td>
<td>.05</td>
<td>20</td>
</tr>
<tr>
<td>Manageability</td>
<td>3.28</td>
<td>3.13</td>
<td>.50***</td>
<td>123</td>
</tr>
<tr>
<td>Overall Assess</td>
<td>3.88</td>
<td>4.23</td>
<td>.53***</td>
<td>&gt;150</td>
</tr>
</tbody>
</table>

*** p < .001, ** p < .01, * p < .05

Notes:

1. The values for the ratings of the 30 competencies are as follows: 4 = "clear strength, exceeds expectations", 3 = "the acceptable standard for students doing well", 2 = "recommended goal for next teaching experience", 1 = "essential goal for next teaching experience".

2) The values for the rating of how difficult the student would "usually find the "management of this/these classes" are as follows: 4="straightforward", 3="reasonable", 2="difficult", 1="very difficult"

3) The values for the rating of "Overall Assessment of Teaching Experience" are as follows:

5= "exceeds expectations", 4 = "competent", 3 = "competent in some areas..."
limitations”, 2 = "limitations across all activities”, 1 = "suitability for teaching questioned”.

### Table 2. Differences between Visiting Lecturer’s and Associate Teacher’s Ratings of "Overall Assessment of Teaching Experience"

<table>
<thead>
<tr>
<th>Subject area</th>
<th>Direction of difference</th>
<th>Lecturer rating 2 points lower</th>
<th>Lecturer rating 1 point lower</th>
<th>Lecturer &amp; Associate T’s rating identical</th>
<th>Lecturer rating 1 point higher</th>
</tr>
</thead>
<tbody>
<tr>
<td>Total</td>
<td></td>
<td>7 (4%)</td>
<td>55 (37%)</td>
<td>74 (50%)</td>
<td>14 (9%)</td>
</tr>
<tr>
<td>N = 150</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Note: The scale values for the ratings are as follows: 5 = "very capable...exceeds expectations", 4 = "competent", 3 = "competent in some areas...some limitations", 2 = "limitations across all activities", 1 = "suitability for teaching questioned".