

# AN ESSAY ON ASSESSMENT: THE ENGLISH CASEBOOK

**TERRY LOCKE**

*Department of Arts and Language Education  
University of Waikato*

**ABSTRACT** *In the late 1980s, there was a clear mandate for change in New Zealand assessment practice in the senior secondary school. However, the establishment of the National Qualification Framework introduced the principle of a single form of assessment for all areas of post-compulsory learning. Like Bursary, the unit standards model of assessment is unsatisfactory. However, there are alternative models of assessment which should be explored as educators mull over their responses to the Green Paper on the National Qualifications Framework.*

## INTRODUCTION

My first audience for this piece of writing is myself. I am doing what I tell my students will work for them: write to clarify. I want to clarify my thoughts about assessment. I want to get a perspective on this issue. I want to do this so that I can reinforce or modify a stance I am taking up with respect to unit standards in English. My major object is to safeguard aspects of English as a subject which I consider to be under threat. In particular, I am thinking about English as:

- driven by teachers driven passionately by a love for language;
- based on a view of language development as holistic and spiral (or recursive);
- drawing on such intellectual traditions as rhetoric and critical literacy;
- wanting students to engage with texts in contexts which encourage them to address the world outside the classroom;
- wanting to extend the competencies of all students; and
- more open-ended and encouraging of inquiry than prescriptive and encouraging check-list teaching.

My second audience consists of my colleagues who are either English teachers or otherwise engaged in the great enterprise of education. What disturbs me in the current situation is the sense of a prevailing orthodoxy which appears to be dampening debate rather than furthering it. What parades as consultation is, in fact, consultation after the fact. That is, it is not really debate at all.<sup>1</sup> I would argue (and in fact have argued elsewhere)<sup>2</sup> that such a state of affairs has been in place since the terms of reference for the development of *English in the New Zealand Curriculum* (henceforth abbreviated to ENZC) were put in place in the early 1990s.

My aim here is to take four approaches to the structuring and assessing of English in the senior secondary school. For each of these, I have proposed seven criteria against which I believe an assessment system itself can usefully be

assessed. Woven into this analysis are references to two narratives. The first of these tells the story of advocacy for change in curriculum and assessment in New Zealand - and what happened to it. The second narrative is necessarily personal - it is my own story as an English teacher, writer and former Head of Department (HoD) English in a large, multi-cultural, co-ed school who has been both actor and acted-on.

## **BURSARY**

Let's begin by acknowledging something rather crucial about Bursary English: its simplicity. Here is a course with four prescribed course components and one assessment component - an externally set, marked and moderated three-hour examination. On the face of it, such simplicity should have a lot going for it. What follows is my 'report' for Bursary in terms of the seven criteria mentioned above.

1. **Scope:** *Does bursary adequately cover the course it is designed to assess?*  
Despite New Zealand Qualifications Authority (NZQA) tinkering, the Bursary English examination is a manifestly inadequate vehicle for assessing the content of the ENZC. Most clearly, it is inadequate as a vehicle for assessing writing and oral skills.
2. **Validity:** *Does it assess what it says it is assessing?*  
It is arguable that it seldom assesses what it says it is assessing. For example, the so-called research essay does not guarantee to test research. It may, like its School Certificate counterpart, test a classroom-based, fabricated research exercise. Moreover, the 60% of the examination which might appear to assess the close reading of literary texts, in practice tests expository writing with little guarantee that the content is anything other than a rehearsed set of teachers' notes.
3. **Best practice:** *Does it encourage 'best practice' in English classrooms?*  
There are ways in which the Bursary Examinations militate against 'best practice':
  - Teaching 'to the examination' encourages the omission of teaching focii that ENZC would otherwise seem to encourage. It also encourages the practice of dispensing sets of (teacher or expert) notes as a safeguard against students failing to engage personally with the studied texts.
  - Section A reading comprehension questions tend to be 'closed' rather than open-ended, thus encouraging students in the belief that reading means hunting for evidence to support an interpretation of a text arrived at by somebody else, i.e. an expert or someone in authority.
  - As mentioned above, the examination can encourage pseudo-research.

- Ironically, by actually incorporating certain ENZC terminology, for example, the discredited use of terms such as 'poetic' and 'transaction' to distinguish text-types, the Bursary examination is in danger of *distorting* students' understanding of the nature of language.

*Can the Bursary Examination support 'best practice'?* Well:

- By default, it at least allows room in the teaching programme for teachers of senior English classes to do interesting things.
  - Its broad-banded and qualitative essay-marking schedules, broad-banded and qualitative as they are, allow student summative achievement to be categorised while recognising that few candidates ever match a category perfectly and without being hooked into the prescription of narrow and artificial, quantified, performance outcomes. Moreover, they allow for excellence.<sup>3</sup>
4. **Assessment mix:** *Is it clearly distinguishable from on-going formative assessment practices?*  
For all its defects, I would argue, the Bursary English examination leaves the fine detail of diagnostic and formative assessment to the classroom - where English teachers, in their professional role, deal with the plethora of knowledge, skills and attitudes that goes into the production of a literate reader, writer, listener, viewer and speaker. For all its shortcomings, the Bursary examination is a clearly defined, summative assessment device.
  5. **Manageability:** *Does it occupy a reasonable percentage of teacher planning time?*  
Yes.
  6. **Moderation:** *Is consistent and equitable moderation practicable, cost-effective and manageable?*  
Yes, given that consistent and equitable moderation is, let us acknowledge it, an impossible ideal. For a start, the assessment tasks are reduced to one event. Secondly, issues of range and level of textual sophistication are dealt to by the nature of the examination and by the extent to which the informal canon of acceptable texts for study has been set in place by decades of teaching practice and by English HODs habitually buying class-sets of the books they studied at university.
  7. **Reporting:** *Does the Bursary English examination lend itself to effective reporting on a student's ability in English?*  
No. On the contrary, once a result has been through the scaling process, it gives little indication of what a student can do in English, particularly with respect to the aims of the English curriculum. Critics of examinations are quick to seize upon their norm-referenced nature. 'In norm-referenced assessment, the achievement of a student is compared with the achievements of a specified group of other students ... Norms describe average or typical

performance, and should not be regarded as standards or desirable levels of attainment.<sup>4</sup> In Bursary English, the more a mark is scaled, that is, brought into line with the norm, the less it becomes indicative of a student's attainment level, even were the raw score itself a valid indicator. I will return to this matter of norm-referencing in my discussion of unit standards.

In terms of these seven criteria, Bursary hasn't scored very well. Not all of the defects I have listed, and readers of this essay may well point to others, are endemic to the examination itself. However, there are enough for us to want to consider other assessment options.

## ACHIEVEMENT-BASED ASSESSMENT

Between 1987 and 1989, the Department of Education, in response to *Learning and Achieving*,<sup>5</sup> and seeking to address the moderation problem in Sixth-Form Certificate, began a series of trials of Standards-Based Assessment. As a member of the Northcote College English Department, I was fortunate enough to be involved in the trials as both teacher and moderator. I say fortunate, because the trials were an excellent opportunity for English teachers to think through a range of assessment issues and, more importantly, to think through systematically what it was their students should be learning in English classrooms.

As it was explained to us at the time, Standards-Based Assessment came in two flavours. The first, Competence-Based Assessment, was related to tasks that you either could or could not do, what we used to call Mastery Learning. In English, the ability to use apostrophes correctly might be termed a competence. The second, Achievement-Based Assessment, was portrayed as the more apt assessment vehicle for subjects like English, where achievement in the various 'aspects' of the subject was deemed to be a matter of degree. In no time at all, teachers were talking fluently in acronyms - practising ABA by assessing students' achievement in terms of the sets of GRCs (Grade-Related Criteria) that had been specially formulated for the subject aspects into which English had been divided.

I am writing here as an ex-HOD who sees a lot of merit in ABA, who introduced it at all levels in Pakuranga College when I was there and who ran many professional development courses, for Longmans, for example, on ABA. How, then, does ABA stack up as an assessment option?

1. **Scope:** *Does it adequately cover the course it is designed to assess?*

At Northcote College, and later at Pakuranga College, I was a member of an English department that successfully used ABA to assess all aspects of English for which those departments wanted summative assessment. In departments, it forced us to address such things as the number of summative assessment points required in any particular aspect and to realise that for summative purposes, we probably tended to over-assess.

2. **Validity:** *Does it assess what it says it is assessing?*

Well, yes, but ....What was really interesting about the introduction of ABA was that it began to open up debate about such things as reading. Teachers who were assessing reading had to face fairly and squarely up to the

question: What do students need to do in order to be able to read? Or more simply: What *is* reading? In 1989, reading was referred to as 'Reception', but that term did not last. In our own department, in order to accommodate two distinct sets of questions which one might use to address a text in the act of reading, we introduced a distinction between 'Content' and 'Genre' for reading and writing (see Figure 1).

### Reading Written, Spoken and Visual Texts

#### A Content

Level 1	Level 2	Level 3	Level 4	Level 5
Can identify straightforward messages of a text.	Can discuss a number of messages of a text.	Can discriminate with respect to the messages of a text.	Can discuss convincingly the underlying assumptions and intentions in the messages of a text.	Can discuss critically the underlying assumptions and intentions in the messages of a text.

#### B Content

Level 1	Level 2	Level 3	Level 4	Level 5
Can identify simple generic features of a text.	Can discuss simple generic features of a text.	Can discuss some part some generic features play in the act of reading a text for meaning.	Can explain the part generic features play in the act of reading a text for meaning.	Convincingly explains and evaluates the part generic features play in the act of reading a text for meaning.

### Writing, Speaking and Image Production

#### Content

Level 1	Level 2	Level 3	Level 4	Level 5
Incorporates straightforward messages in the production of a text.	Incorporates a number of linked ideas in the production of a text.	Incorporates and develops a number of ideas in the production of a text.	Incorporates and consciously develops a number of ideas in the production of a text.	Incorporates and consciously develops with imagination and flair a number of ideas in the production of a text.

#### Style

Level 1	Level 2	Level 3	Level 4	Level 5
Can employ some of the features of a genre in a straightforward way.	Can employ a range of features of a genre.	Can employ a range of features of a genre for deliberate effect.	Shows confidence in employing a range of generic features for deliberate effect.	Shows confidence and originality in employing a range of generic features for deliberate and striking effect.

Figure 1: Assessment levels

3. **Best practice:** Does it encourage 'best practice' in English classrooms? In a number of ways, ABA enhanced 'best practice':

- Students began to be given some clear directions as to what grades actually meant and what had to be done to achieve them.
- Teachers began writing assessment schedules or rubrics for specific tasks which attempted to spell out in detail what a particular achievement level meant for the given task. Inevitably, teachers and students had to begin paying detailed attention to a whole range of features of a range of texts, and an assessment schedule is shown in Figure 2. A number of points might be made with respect to this resource. Firstly, it is starting to resemble, particularly in its qualitative aspect, the broad-banded essay-marking schedule of a Bursary English examination.

Secondly, the more prescriptive it becomes, the more contentious it has the potential of becoming. Under 'Content', for example, it would appear to suggest that stories should have complex characters. Now that is debatable. Some narrative genres lend themselves to stereotypical characterisation. If we insist on the prescription of the marking schedule, we are in danger of being lured into bad classroom practices, a point I will make later with respect to Unit Standards.

- In major aspects of English such as reading and writing, students were engaged in widely-spaced assessment events with a range of text-types or genres. Their final grade was generally assigned by the teacher as a typicality of the grades achieved for each of these discrete events. Typicality was the term used to denote the level of performance which typified a student's performance in a particular aspect of English such as close reading. Consequently, students had an incentive to lift their performance, knowing that an improvement would gain recognition in their final grade.

## **Content**

**Level 1:** *Straightforward ideas.*

The story tends to revolve around one or two simple events. There is little sense of a theme or idea which the story might serve to illustrate. Characters are straightforward and only sketchily differentiated. Links between cause and effect may be rather tenuous.

## **Genre**

**Level 1:** *Employs some of the features in a straightforward way.*

The plot consists of a series of simple events rather casually linked. Pace will be inconsistent. A setting is suggested but has little relevance to the action. Narrative point of view may show some inconsistency and difficulty with tense. Characters tend to be defined by what they do. Characterisation is limited.

**Level 2:** *A number of linked ideas.*

The writer begins to show an awareness of how a story can serve to illustrate an idea or theme. Characters are clearly distinguished one from the other and appear to act out of some motive. Events are linked plausibly with one another.

**Level 3.** *Develops a number of ideas.*

There is a clear sense that the narrative is serving to illustrate one or more ideas or themes. Characters are clearly defined, act out of a motive and are somewhat complex (i.e. capable of experiencing internal conflict). Events are convincingly linked.

**Level 4:** *Consciously develops a number of ideas.*

The narrative convincingly encapsulates one or more ideas or themes which serve to underpin and unify it. A range of complex characters are deliberately evoked to act as a vehicle for the action and to highlight aspects of the theme. There is evidence that the writer has thought about the cultural setting of their narrative and the disposition of their intended audience.

**Level 2:** *Employs a range of features.*

The plot is characterised by a clear, linear sequence of cause and effect. There may be some inconsistencies in pace. The writer has selected and described an appropriate setting. A consistent point of view is maintained with perhaps a few lapses. Characters tend to be defined by how they think as well as by what they do. Characterisation is sound.

**Level 3.** *Employs a range of features for deliberate effect.*

Plot is characterised by conscious planning, e.g. the incorporation of such devices as predicament, complication, resolution, key choice within a chronological sequence. The story is consistently paced. Setting is appropriate and clearly evoked. Point of view is consistent and chosen for deliberate effect. Characters deliberately contrast each other and are evoked through a variety of means. There is evidence of a deliberate attention to "literary" style (in choice of diction and syntax).

**Level 4:** *Confidently employs a range of features for deliberate effect.*

The plot is consciously structured and unified by the placement of such features as a climax, turning point or denouement within a chronological sequence. An appropriate setting has been carefully chosen and both directly and indirectly evoked (e.g. through dialogue). Point of view is confidently handled and sustained. Style is confident, inventive, fluent and shows a wide vocabulary.

**Level 5:** *Consciously develops with imagination and flair a number of ideas.*

The narrative convincingly, subtly and cleverly encapsulates one or more ideas of themes which serve to underpin and unify it. A range of complex and appropriately contrasting characters are deliberately evoked to act as a vehicle for the action and to highlight aspects of the theme. There is evidence that the writer has thought carefully about and even researched the cultural setting of their narrative and the disposition of their intended audience.

**Level 5:** *Shows confidence and originality in employing a range of features for deliberate and striking effect.*

The plot is consciously unified and shows a willingness to restructure the chronological flow of events in the story through the confident use of either a single point of view or of multiple perspectives on the action. Settings are cunningly evoked and illuminate the action. The writer's style is individual, inventive and innovative with a wide-ranging vocabulary and a command of a variety of syntactical structures.

Figure 2: Form 6 English: Writing Literary Narrative: Marking Schedule

4. **Assessment mix:** *Is it clearly distinguishable from on-going formative assessment practices?*

Because ABA worked in terms of a set of broad, or global, grade-related criteria, it was clearly a summative tool. Summative assessment points were defined at the start of the year. Different aspects, what we would now call the strands, of English were given weightings relative to their perceived importance. I know no school that weighted Oral, Written and Visual Language equally, as Unit Standards has done. Some aspects might have only one summative assessment, for example, Listening, Processing Information, but teachers would be clear that all summative assessment tasks would be prepared for by means of the ongoing formative assessment that took place in the classroom. Re-testing was not really an issue, because students had prior experience of evaluation, either through this formative assessment or through a prior summative assessment event.

5. **Manageability:** *Does it occupy a reasonable percentage of teacher planning time?*

Unlike the Bursary system, HODs and individual English teachers were involved in the writing of specific marking schedules for assessment events tailored to the needs and learning activities of individual classes. In the case of Pakuranga College, the global GRCs developed by the Department of Education during the trials were also re-examined and rewritten.

Such a process *is* time-consuming. But the spin-off from the demand for consultation and cooperation was debate and discussion of issues that for years had been swept under the carpet. Moreover, under ABA, assessment became far more integrated into classroom practice, since teachers began relating lesson-specific learning objectives to the global GRCs their English departments were working with, and began planning learning activities that had a clear relationship to their lesson or module objectives.

6. **Moderation:** *Is consistent and equitable moderation practicable, cost-effective and manageable?*

At the school level, moderation involved a range of processes such as common tests, check-marking, syndicalisation, team-marking, the selection of exemplars, and, above all, the sharing of expertise through joint planning. In my own experience, with a relatively stable English department, the goal of consistent and equitable moderation became relatively easily attained.

At a national level, as with any system involving internal assessment, moderation becomes a more complicated and costly process. Intra-departmental consistency is more achievable in a school than inter-departmental consistency across a number of schools. A number of factors potentially affect the attainment of consistency. These include: the range of texts demanded, the level of sophistication of texts to be engaged with, the finer points of interpreting GRCs (at the level of global criteria as well as in their translation into specific marking schedules), the setting of assessment tasks and in the marking of those tasks.

However, as I will argue later, these difficulties deserve to be grappled with, even in the knowledge that a totally consistent and equitable system of moderation is both an impossible ideal and that agreement is unlikely ever to be reached by all parties on how such moderation might best be implemented.

7. **Reporting.** *Does the ABA lend itself to effective reporting on a student's ability in English?*

At Pakuranga College, English students were given course statements at the beginning of the year which included the grade-related criteria that would be used for assessment purposes. They were also told how many assessments would occur for each aspect of English. They received reports twice a year. Each report contained a brief statement detailing what each assessment aspect was about. Students were given a grade typical of their performance in each aspect of English at the time of reporting. While expressed numerically, the grade corresponded to a written descriptor related to an aspect of English which the report explained to its reading audience.

### **What happened to ABA?**

As I see it, Achievement-Based Assessment had and still has a lot going for it. It allows for an adequate scope in the assessment of English and, if couched in the right terms, can assess what it says it is assessing. Furthermore, it can facilitate good classroom practices and has built into it an emphasis on formative assessment. Although it puts a number of time and professional demands on English teachers, I would say it enhances their professionalism. But it raises all sorts of questions about moderation, especially at a national level.

As it turned out, ABA never made it as a serious contender for seventh-form, national assessment honours. In 1988, the Hawke Report<sup>6</sup> signalled the government's intention to establish a standards-based qualifications framework that would comprehensively reshape post-compulsory qualifications in this country. Standards-based assessment was to become the norm for achievement in

all areas of learning, but in a shape that would be markedly different from the achievement-based assessment many English teachers had become used to working with. That shape and the advocacy for it will be examined later in this essay. But first, I want to engage in a diversion with respect to *English in the New Zealand Curriculum*.

### **English in the New Zealand Curriculum**

At the end of 1988, the English Department at Northcote College asked a rather obvious question. If ABA was proving such an excellent system of assessment in the senior secondary school, why not draft a set of grade-related criteria (GRCs) for use in the school's junior secondary English programme? Which is what we did. A set of aspects was decided on and sets of GRCs (with four rather than five levels) were drafted and implemented in 1990.

What was interesting about these junior GRCs was their remarkable similarity to their senior counterparts. I can distinctly remember the question being asked of us at meetings where we explained this new system to parents. How was it that we were assessing students in Forms 3, 4 and 5 according to the same criteria? At first, we tried suggesting that, in fact, many students would not get to Level 4 (the highest grade) until Form 4 or Form 5. But, in fact, when we began using these criteria in our classrooms we realised that able students were reaching Level 4 in the third form, albeit with texts that were less challenging than those being used by fifth-formers.

This experiment had inadvertently thrown up a singular truth: when it comes to literacy skills, the tasks readers and writers need to perform do not change much with age. What changes are such factors as the level of textual sophistication, the range of texts they are dealing with and the social contexts within which their performance as readers and writers is being put to the test.

Because I have dealt extensively with ENZC elsewhere, I will just briefly list its flaws here, but note that these are crucial to the argument against Unit Standards that follows.

- **Levels:** The enforced commitment to eight levels of achievement imposed on the developers of ENZC was a serious mistake. The eight levels were arbitrary and had no basis in the current understanding of literacy(ies) development. Moreover, they ran counter to ENZC's insistence that literacy development is holistic and recursive. Moreover, the levels generated sets of achievement objectives which were distinguished by faulty logic and language which confused readers the more they tried to 'unpack' them. Finally, the levels suggested an age-equivalence which poses the serious danger of our students being 'dumbed down', particularly at primary, intermediate and lower secondary school, when teachers use the ENZC achievement levels to guide them in setting learning objectives for their students.<sup>7</sup>
- **Terminology:** ENZC used terminology for distinguishing text categories, that is, 'expressive', 'poetic' and 'transactional' that dates back to British thinking about language development in the late 60s and early 70s and

which has been discredited for at least ten years. The persistence with this terminology actually has the potential to distort students' understanding of language function. Moreover, the term 'function', used in the document in contradistinction to language 'process' is both confusing and belies its usual usage in treatments on the nature of language.

The consensus of opinion that responded to the draft *English in the New Zealand Curriculum* overwhelmingly rejected the levels but was ignored by the Ministry of Education.

### **And So To Unit Standards**

In a carefully argued and documented paper entitled *Advocacy, Evolution and Learning: School Assessment for the National Qualifications Framework* (1995)<sup>8</sup>, Bill Lennox argues "...that the National Qualifications Framework assessment model is consistent with the advocacy for change and the evolution of assessment systems for school qualifications."<sup>9</sup> To establish his case, Lennox examines a range of literature on the theme of assessment dating from the 1967 collection, *Educating New Zealanders* (Jack Shallcrass). Lennox's paper is a reading of the documentation which seeks to find what Lennox *wants* to find, that is, an implicit advocacy for the National Qualifications Framework's use of 'competency-based' assessment, what he would prefer to call 'single-standard' assessment. I say implicit, because Lennox himself has to concede that:

The single-standard form of assessment introduced by the Qualifications Framework is not consistent with the advocacy, which generally presumed graded criteria. Assessment against single standards was not advocated (but see 10 below).

Lennox's 10 below reads:

Although the Framework's single standard concept was not consciously advocated, it was practised.<sup>10</sup>

Lennox supports the latter statement by an analysis of both Bursary and ABA assessment practices which, I would contend, are highly questionable. I offer the following as examples of the kind of argument being offered.

If a researcher were to monitor the meetings and work of English markers there is little doubt that a sense of an attainable single standard essay would emerge, perhaps even quite explicitly, and that it would be a B essay.<sup>11</sup>

Teachers are currently using standards-based criteria that often specify single preferred and attainable standards. Teachers have tended to the level 4 performance as the highest feasible goal for all students.<sup>12</sup>

As I will be arguing in the next section, such claims are not only dubious. They simplify the context of the assessment situation. An essay written in a Bursary English examination is a far more 'contained' event than the writing of a genre of a student's choosing within the context of an internally assessed environment. They also beg certain questions, such as, Is assessment about identifying the highest feasible goal for all students?

On the basis of my reading of Lennox's valuable paper, the clearest reflection of the assessment advocacy the writer reviews is the *Learning and Achieving* report of June, 1986. As Lennox summarises it, the committee recommended that:

For form 5:

- schools should be able to offer courses of their own design
- assessment should be wholly school based
- assessment should relate to course completion and to standards expressed at three levels
- a descriptive profile-style fifth form certificate should replace School Certificate

For form 6

- schools should be able to offer courses of their own design
- course length should be variable
- students should be able to study at a variety of levels, part-time and at institutions other than schools
- assessment should continue to be wholly school based
- assessment should be achievement based against criteria at four levels
- national moderation to ensure comparability of assessment within subjects should be introduced to suit each subject and not tied to results at other levels.

For form 7

- the curriculum should be achievement based
- assessment should be by national examination and school based
- assessment should be achievement based related to five levels of achievement
- one seventh form certificate would report school and examination results separately
- a merit award would be available and entrance to university would be by a combination of school and examination results.

I would contend that these recommendations more authentically reflect the advocacy which unit standards and the National Framework claim to represent but in fact depart from.

## UNIT STANDARDS

Currently up to 150 schools are trialling English unit standards. In addition to their unit standards packs (containing unit standards descriptions), teachers have

received (as of August, 1997) 17 newsletters to incorporate into their Assessment Guide totalling nearly 600 pages of printed material.

As mentioned previously, unit standards are couched in terms of single defined standards or competencies. 'Credit is awarded when the assessor is confident that the student has performed to the required standard, and that this performance could be repeated.'<sup>13</sup> There are 17 Level 1 unit standards in English, 11 at Level 2, 8 at Level 3 and 4 at Level 4. Unit standards are not course-related. Theoretically, a student in a 'Form 6' class could be assessed at any of Levels 1, 2, 3 or 4.

In practice, each of the English unit standards is broken down into two to four elements. Each of these elements has attached to it a number of 'performance criteria'. 'The standard of performance required of a student to achieve credit is set initially in the performance criteria of the unit standards'.<sup>14</sup> A teacher of a sixth-form English class could conceivably be applying 92 elements with a total of 231 performance criteria to the assessment of the students.

Moreover, in designing assessment tasks to fit these performance criteria, teachers need to take on board a series of prescribed range constraints. Range statements are used to further prescribe the parameters of a student's required performance. For example, Element 1 of Unit Standard 8820 (See below: ENGLISH WRITTEN LANGUAGE: Read written texts closely) is 'Read poetic text closely' and has three performance criteria. The second of these (1.2) specifies the outcome: 'Language features are analysed in terms of their effects'. This outcome is constrained by a range statement which says that the analysis '...must include at least three language features from each of the following categories - figures of speech, choice of words, grammar'. As *English Trials Newsletter 13* shows, students cannot be deemed as having fulfilled this performance criterion unless they in fact analyse three of each of the specified features.<sup>15</sup>

Comparable elements from the unit standards in English have been reproduced below. I will be contending that these unit standards typically embody the flaws that beset this whole enterprise. Having previously applied my seven criteria to Bursary English and to Achievement-Based Assessment, I will now apply them to unit standards.

8820: ENGLISH WRITTEN LANGUAGE

Read written texts closely

*element 1: Read poetic text closely.*

*performance criteria*

- 1.1. Meanings of the text are explained in terms of its ideas.
- 1.2. Language features are analysed in terms of their effects.  
(Range: must include at least three language features from each of the following categories - figures of speech, choice of words, grammar.)

- 1.3. Ideas in the text are critically evaluated in terms of the historical or social context of the text.  
(Range: must include at least two ideas.)

8831: ENGLISH WRITTEN LANGUAGE

Read closely and evaluate written text

*element 1: Read closely and evaluate the effectiveness of poetic texts.*

*performance criteria*

- 1.1. Explanation of meanings of the text identifies its ideas.
  - 1.2. Evaluation of language features identifies their effectiveness.  
(Range: must include at least three language features chosen from each of the following categories - figures of speech, choice of words, grammar.)
  - 1.3. Critical evaluation of ideas in the text identifies the text's historical or social context.  
(Range: must include at least three ideas.)
1. **Scope:** *Does the unit standards system adequately cover the course it is designed to assess?*  
Overall, representations of the unit standards matrix for English indicate a determination on the writers' part to provide a reasonably full coverage of the English curriculum document. (As I will be suggesting later, however, there are disadvantages in sticking to the letter of ENZC.) If we disregard unit standards at Level 4, we find that there are gaps with respect to 'Processing Information' at Level 3 (a fault, in my view), gaps with respect to 'Expressive Writing at Levels 2 and 3, and a gap with respect to 'Interpersonal Speaking and Listening' at Level 3.

2. **Validity:** *Does the system assess what it says it is assessing?*  
This is a crucial question, especially given the Framework's vaunted claim as presenting clear, attainable, measurable learning outcomes. Unit standard 8820, reproduced above, is designed to assess students' ability to 'Read poetic text closely'. If we put to one side the use of the term 'poetic' for a moment, we can all agree that we want students to read closely.

In terms of this model, the performance criteria should tell us what it means to read closely. In the first place, it is suggested, if 'Meanings of the text are explained in terms of its ideas,' then we have evidence of close reading. The first task of students and teachers, of course, is to understand what this criterion means. To this reader, it appears to suggest that the 'meanings' of a text can be accounted for by the 'ideas' in the text. Is this circular reasoning? Is it nonsense? I would argue that when we read, we actively construct meanings as a result of the way we question texts. I would

have to wonder whether this criterion actually assesses reading as I understand it.

The second criterion would appear to suggest that close reading involves a determination to give equal weight to the analysis of figures of speech, diction and grammar. Again, I would argue that this is not close reading as I understand it. A close reader, in my understanding, is one who can discriminate *between* linguistic features of a text in terms of their relative importance in impacting upon a reading. Rather, this criterion would appear to be assessing the ability to work through a text with a checklist of linguistic features. This also is not close reading.

The third criterion suggests that close reading involves a reader in critically evaluating the 'ideas in the text' in terms of 'the historical or social context of the text'. As I read this criterion, it is telling students that critical reading requires them to somehow adopt the cultural perspective of the era out of which a text has been produced. Now that is not what I understand by critical reading either. And it is certainly remote from my understanding of critical literacy. What I would want from my students, is the ability to identify the ideological positions implicit in a text and to critique them in terms of their own developing stance. That is what I would call critical literacy. But that is not what this criterion is assessing.

In an article in *English in Aotearoa*, Warwick Elley dismisses the English unit standards because 'They are too fuzzy to clarify what students know and can do'.<sup>16</sup> On the basis of this analysis, I have to agree with him.

The other central problem with unit standards is their suggestion that single-standard performance criteria can be used to differentiate between students on the basis of *levels* of performance. For levels to have meaning, an observer reading unit standards 8820 and 8831 should readily perceive the difference in performance expected. What *is* the difference? There appears to be little differentiation between performance criterion 1.1 in each. They are both equally confusing. There also appears to be no difference between performance criterion 1.3. The one difference is in 1.2 where at the higher level (Form 7), students are expected to evaluate rather than analyse a rigidly constrained set of language features. The differentiation depends on an assumption that analysis precludes evaluation - a dubious assumption, I would contend.

There is a simple explanation for the highly questionable differentiation being used here. The writers of the English unit standards were drawing heavily on ENZC. Now, as argued previously, ENZC is flawed by its adoption of a system of levels that defies everything we know about language learning. As Elley writes:

... English is one of those subjects that does not lend itself to standards-based assessment. This kind of assessment can only work if the standards are clearly stated, which they are not, and if there is a clear progression from lower to higher standards, which there is not. In English, most of the key objectives are very similar from year to year. They nearly all revolve around *learning to communicate competently in spoken and written forms*. These aims are

general, rather than specific, and any progressions are found not in different skills learned each year, but rather in an increasing level of sophistication in the texts that students are expected to comprehend and generate.<sup>17</sup>

ECNZ, in its 'Close Reading' achievement objectives, does not introduce 'evaluation' as a differentiating factor until Level 8. Hence, the writers of the English unit standards do not introduce 'evaluation' until Level 3. But as Elley argues, such activities as interpretation, analysis and evaluation are key objectives in any English programme, including primary school language programmes.

Their use as factors differentiating between unit standards levels is spurious. And, as I will go on to suggest, such usage is also damaging to 'best practice' in English classrooms.

3. **Best practice.** *Does the system encourage 'best practice' in English classrooms?*  
There are a number of ways in which the unit standards system is damaging to best English classroom practice. I will mention but a few here.

- As argued above, the factors used to differentiate between the unit standard levels are based on a flawed concept of literacy development. If English teachers incorporate these factors into their programme planning, they will be building in a flawed understanding of literacy - and passing that on to their students.
- As argued previously in this essay and elsewhere, ENZC's use of such terminology as 'poetic', 'transactional', 'expressive', 'function' and 'process' is problematical because their use militates against an adequate understanding of such crucial factors as language function and the social nature of genres. I keep seeing evidence of students' language understanding being distorted by the way these terms are used and it is frankly disturbing.
- As my analysis of the performance criteria for unit standard 8820 indicates, the formulation of its performance criteria distorts the nature of the reading process in two ways, firstly through a kind of misleading vagueness and secondly through a dubious kind of prescriptiveness. I'm inclined to see the former as the less damaging of the two. The problem with the second is that it distorts students' understanding of the role of stylistic features in impacting upon the construction of a reading. Good readers can discriminate between the relative importance of particular stylistic features. They don't read with a check-list. Moreover, the distinction between 'figures of speech' and 'choice of words' is arbitrary and misleading.

I must mention a story I was told of a HOD English who spent an afternoon looking for poems to fit the performance criteria! What message is this communicating about poetry? Does a poem need at least three figures of speech before it is a poem? What about prose writers like Hemingway and Sargeson who

studiously downplayed the use of metaphor? Are they to be exiled from the unit standards English canon?

There is a central difficulty with range statements of the type I am addressing here. The unit standards system would appear to demand them because of the enormous moderation difficulties the system poses. But, the more specific the range statements, I would argue, the more they militate against sound teaching and classroom practice.

Finally, unit standards with their emphasis on single-standard outcomes play into the hands of those who would seek to move the curriculum in the direction of vocational education with its emphasis on discrete, transportable skills, which, according to Michael Peters and James Marshall, are far easier to measure than 'knowledge' and 'understanding'.<sup>18</sup> With respect to English, I would argue that such an emphasis has the potential to stultify, for example, the further exploration of the implications for classroom practice of such exciting perspectives as 'Critical Literacy' with its emphasis on the development of a critical understanding of how texts operate in the world.

4. **Assessment mix:** *Is it clearly distinguishable from on-going formative assessment practices?*

My reading of the trials newsletters suggests to me that, as always, when English teachers get together, a lot of valuable exchanges occur related to helping students become better readers and writers. I have been suggesting that the unit standards system is flawed. Part of the problem, I would suggest, is that an effective assessment system has to assess all elements of the curriculum in a summative sense. However, the more summative tests, especially in an internally assessed system, the more difficult is the process of moderation. I would suggest that the unit standards system, in failing to distinguish between formative and summative assessment, has made a rod for its own back and for the backs of teachers. A better system would focus on ways in which formative assessment might take place, by the development of freely available rubrics, sets of criteria and, most importantly, learning materials, and reduce summative assessment to a bare minimum.

5. **Manageability:** *Does the system occupy a reasonable percentage of teacher planning time?*

No.

6. **Moderation:** *Is consistent and equitable moderation practicable, cost-effective and manageable?*

No, for a number of reasons already discussed.

- Inter-level differentiation is flawed.
- Performance criteria are faulty.
- There are too many summative tests and performance criteria.

7. **Reporting:** *Does the unit standards system lend itself to effective reporting on a student's ability in English?*

This brings us to the final difficulty with unit standards. As Elley argues, unit standards '... are not standards at all .... They do not differentiate between students of widely differing levels of achievement'.<sup>19</sup> What becomes clear to anyone reading unit standard 8820 is that, even without the shortcomings analysed above, it simply does not tell anyone *how well a student reads*. Ironically, such a unit standard credit in English will tell an employer little more than the mark in the English Bursary examination, which the unit standards system is intended to be such an advance on.

There is an irony in all of this. Proponents of unit standards have been quick to present them as an escape from norm-referencing. But are they? In Lennox's paper, which I have quoted from in a number of instances, he draws a conclusion that: 'The Framework's form of a single attainable standard is present to an increasing degree in some recent examination marking schedules. There was a clear trend within marking schedules towards the identification of a single standard, usually at the second to top category [B], that is attainable by a large proportion of learners'.<sup>20</sup>

One way of thinking about unit standards is as an attempt to describe a typical performance worthy of being given a pass in an aspect of a subject. But is not this just another form of norm-referencing? John Codd, commenting on NZQA's definition of an 'agreed' standard as 'the level of performance that is considered acceptable ... [and] ... established only in practice, by consultation and agreement among assessors and moderators', has this to say: 'Clearly, this is little more than a linguistic manoeuvre necessitated by the fact that the so-called 'standards-based' assessment upon which unit standards are said to be based cannot avoid 'norm-referenced' judgements'.<sup>21</sup>

### **The Victorian Certificate of Education: Original English Study Design**

'The development of the VCE,' writes Terry Hayes, currently President of the Victorian Association for the Teaching of English, 'was an attempt to bring some form of symmetry and order to an educational scene characterised by variety and the rapid proliferation of courses across all learning areas'. He goes on to assert that: 'The design was premised on the fact that English was essentially an agent of community and an English study design should enable students to develop a capacity to be critically literate in an increasingly diverse range of communication contexts while, at the same time, recognise the cultural and social dimensions in which they developed their abilities. Essentially it was designed to make a potential 55,000 students (the average VCE cohort over a two-year cycle) feel that English was doable, purposeful, relevant, and, in appropriate contexts, collaborative'.<sup>22</sup>

In its original form, the VCE study design required students in the last two years of secondary school to take four units of English. Units 1 and 2 would constitute what we might call a Form 6 course. Units 3 and 4 constitute an equivalent to our Form 7 course.

Of interest to this writer is the flexibility the study design offers schools for assessing and reporting on student work in the Form 6 year: 'To satisfactorily complete a unit of study, the student must satisfactorily complete all the work requirements for the unit in accordance with the specifications in this study design. Assessment of satisfactory completion will be made by the schools. Schools will report the student's result for each unit to VCAB [Victorian Curriculum and Assessment Board] as S (Satisfactory) or N (Not Satisfactory)'. The student's completion or non-completion is reported on the Statement of Results issued by the VCAB. 'Procedures for assessment of levels of performance in Units 1 and 2 are a matter for school decision,' but are not reported by the VCAB.<sup>23</sup>

For Units 3 and 4 (our Form 7), the completion requirements which pertain to Units 1 and 2 also apply and are reported on. In addition, however, students are assessed by means of four common assessment tasks (CATs) supervised by the VCAB. More of these later. The student's assessment for each CAT is reported on their Statement of Results, using a ten-point scale using five grades with two levels within each grade: A+, A, B+, C, C+, C, D+, D, E+, E. Each grade is accompanied by a description of the level of performance typically associated with the grade awarded.

The English study design has three areas of study:

1. The Craft of Writing: '... intended to improve competence and confidence in writing for a range of purposes and audiences in a variety of forms. The role of oral language activities in developing writing skills is emphasised'.

2. Reading and the Study of Texts: '... focuses on reading a wide range of texts with comprehension, enjoyment and discrimination. It emphasises a variety of forms of responses to texts: oral as well as written, and creative as well as analytical'.

3. The Presentation of Issues and Argument: '... focuses on the use of language for informative and persuasive purposes. It is intended to develop students' abilities to evaluate information received, think critically, logically and creatively about a variety of issues, and present personal points of view effectively both in speaking and writing'.<sup>24</sup>

Students in each unit are given a set of work requirements which '... provide the framework within which the areas of study are to be explored and the objectives achieved. They draw together work undertaken in each of the areas of study'.<sup>25</sup> These work requirements are summarised below.

**Common requirements for each two-unit sequence:**

1. Workbook: an organised collection of work done in the course.
2. Writing folio: a collection of 'finished' writing for a range of purposes and audiences in a variety of forms.

3. Text response: a collection of 'finished' pieces of work which result from reading, interpreting and responding to a range of texts, in a variety of forms.
4. Communication project: a language-based activity in which students develop and present a product to an audience other than their teacher.

**Work requirements specific to particular units:**

- 5a) Response to issues (Units 1 & 2): a sustained piece of writing or oral presentation.
- 5b) Presentation of an issue (Unit 3): a piece of writing in two sections, analysing and presenting points of view on a topical issue, and an oral presentation of an aspect of the issue.

Each of these work requirements is circumscribed by a set of constraints described more fully in the study design booklet.

Finally, we come to the CATs - the common assessment tasks administered by the VCAB for Form 7 students. I will provide a minimum of description for each common assessment task.

**CAT 1. *Presentation of an issue:*** Students produce a piece of writing in two sections. (Internally assessed but subject to external moderation.)

**CAT 2. *Writing folio:*** Students present a folio of selected pieces of writing. (Internally assessed but subject to external moderation.)

**CAT 3. *Oral communication:*** Students participate in three oral communication activities in at least two different audience settings and for at least two different purposes. (Internally assessed. No external moderation.)

**CAT 4. *Text response:*** Students complete one piece of sustained writing on one of the selected texts. (An external two-hour examination, double marked by a VCAB panel.)

How might the VCE English study design, at least in its original form<sup>26</sup>, stack up against the seven criteria I have used previously as a yardstick for measuring the Bursary English examination, ABA and unit standards?

1. **Scope:** *Does it adequately cover the course it is designed to assess?*  
The rather brief summary of the English study design I have included here doesn't do justice to the range of language experiences embraced by the work requirements of this course. A reading of the work requirements in detail reminds one of the sort of course a progressive and demanding New Zealand English department might offer to its sixth-form students.

2. **Validity:** *Does it assess what it says it is assessing?*  
I would have some reservations about the shape and format of the external examination CAT. On the face of it, it tests writing rather than reading. However, the other CATs closely match the competencies they claim to be assessing.
3. **Best practice:** *Does it encourage 'best practice' in English classrooms?*  
I would argue that the use of broad criteria for assessment within the context of a fairly tightly defined task encourages 'best' classroom practice because it eliminates the deleterious effects of the tightly prescribed (and often badly written) criteria and range statements of the current unit standards trial. Moreover, the emphasis on satisfactory task completion actually encourages a process of negotiation between the teacher/expert and student/apprentice to define when a student has reached an adequate and appropriate standard. Finally, the injurious effects of the structure and terminology of ENZC are avoided in the study design.
4. **Assessment mix:** *Is it clearly distinguishable from on-going formative assessment practices?*  
In a 1996 article, Warwick Elley makes the point that '... good diagnostic assessment is best done informally, according to need, in unstandardised fashion. By contrast, accountability assessment must be uniform and tightly controlled'.<sup>27</sup>

The VCE English study design clearly differentiates between the diagnostic and formative assessment done in the classroom and what Elley calls 'accountability assessment' - summative assessment when it enters the public domain.

What mediates between the formative practice in the classroom and the summative assessment of the CAT are sets of criteria which are, in fact, qualities rather than prescribed outcomes. Here are the criteria used for the assessment of CAT 1: Presentation of an issue.

#### **Thought and content**

(The quality of thinking demonstrated in the work)

- understanding of the demands of both parts of the task
- knowledge and control of the chosen content
- complexity of the ideas and issues explored in the work

#### **Structure and organisation**

(The quality of the structure developed by the student in response to the task as specified)

- the shape and cohesion of the work
- the sequence and coherence of the work as it develops
- striking or imaginative organisation of the material

#### **Expression, style and mechanics**

(The quality of the language used in organising and communicating the issues and ideas presented)

- the expressiveness and fluency of the work
- the effectiveness and appropriateness of the language chosen, given the task
- striking or imaginative use of language
- control of the mechanics of the English language to support meaning.

These qualities are well fitted to shape the sorts of specific objectives English teachers construct in their lesson planning. In the summative assessment process, teachers initially gain an impression of the student's work in terms of the above criteria. Then they assign the student with a grade from A+ to E according to the mark-banded descriptors for the particular CAT. For Presentation of an Issue, these descriptors are as follows:

**A** Achieved a critical and controlled analysis of the language used in the presentation of an issue. Demonstrated a clear understanding of the issue and its wider implications in a substantiated view which was organised and cohesive. Used fluent and persuasive language with assurance and precision.

**B** Achieved a focused exploration and analysis of the language used in the presentation of an issue. Demonstrated understanding of the complexities of the issue. Achieved effective selection and ordering of ideas. Presented a view on the issue with confidence, using fluent and persuasive language.

**C** Achieved some analysis of the language used in the presentation of the issue. Identified major aspects of the issue. Achieved a coherent ordering of ideas and expressed a view in clear and accurate language.

**D** Achieved some discussion of the language used in the presentation of the issue. Identified some aspect of the issue. Presented a clear view on the issue. Demonstrated adequate organisation and language control.

**E** Identified some views presented on an issue. Presented some ideas relevant to the issue. Demonstrated some organisation and language control.

It is worth keeping in mind that each band allows the allocation of either a Grade or a Grade +, thus creating a 10 - point scale. Unsurprisingly, such descriptors remind one both of the mark-banded assessment schedules for Bursary English and of the grade-related criteria of the 1987-9 trials, except for the use of 10 steps instead of 5.

5. **Manageability:** *Does it occupy a reasonable percentage of teacher planning time?* Like ABA and unit standards, and unlike Bursary English, the English study design encourages teachers to think in terms of objectives and criteria. In

terms of workload, I would see it as commensurate with the sorts of changes brought about in departments which adopted ABA. However, the workload pales in comparison with the demands of unit standards.

6. **Moderation:** *Is consistent and equitable moderation practicable, cost-effective and manageable?*

The English study design put a high value on the professionalism of teachers while maintaining control over Elley's 'accountability' assessment. One might say that, at least in its original form, it achieves a rather nice balance. One task is solely internally assessed. One task is solely externally assessed. Two tasks are internally assessed and externally moderated.

No moderation process for internally assessed tasks is cheap or unproblematical. I would argue, however, that the English study design is a fair compromise. Moreover, by limiting and tightly defining (as Elley requires) the number of common assessment tasks, it is eminently manageable.

7. **Reporting:** *Does the English study design lend itself to effective reporting on a student's ability in English?*

As indicated previously, the student's Statement of Results indicates both the student's task completion/non-completion record and their grades for each CAT. Each grade is accompanied by a description of the level of performance typically associated with the grade. A student receiving a B for CAT 1 would have the following description on their Statement of Results: 'Achieved a focused exploration and analysis of the language used in the presentation of an issue. Demonstrated understanding of the complexities of the issue. Achieved effective selection and ordering of ideas. Presented a view on the issue with confidence, using fluent and persuasive language'. We are clearly back in the territory of broad-banded descriptors which, as we have already seen, have been used by Bursary English markers for years. What is different here is that the description actually enters the public domain through the reporting process.

## CONCLUSION

At the time of writing, the Government has just released its Green Paper on the National Qualifications Framework. A cursory examination of its key proposals would indicate that NZQA is adhering to the view that '... all qualifications should have clearly stated *outcomes* [statements about what students know or can do], and be capable of being related to one another through *level* and *credit values*'. At the same time, the NQF is projected as '... able to accommodate all types of qualifications, from school examinations to degrees, whether or not they use unit standards'.<sup>28</sup> Is this really a change of course? Is NZQA really prepared to take on board the criticism of the current assessment practices operating in senior English as discussed in this essay? Does it actually *understand* the criticism? Time will tell.

In the meantime, 112 (at last count) English departments throughout New Zealand are voicing their rejection of, or reservations about, the system of unit

standards. Moreover, the PPTA, in its report *Te Tiro Hou*, has rejected competency-based assessment in English as neither valid, fair, nor motivating. I believe the time has come for us to look beyond Bursary English and unit standards to a system such as the VCE English study design. I believe it presents a balanced way of approaching the teaching and assessment of English in the senior secondary school which avoids the pitfall of a single public examination and the invidious nonsense of unit standards. I believe it accommodates nicely the advances in thinking represented by ENZC without taking on board its flawed framework and terminology. Finally, and ironically, it is almost an exact mirror of the assessment advocacy represented in the old Department of Education's *Learning and Achieving* report of 1986.

### FOOTNOTES

1. I would recommend a careful reading of Ian Frowe (1992). *Persuasive forces: language, ideology and education*. In Richard Andrews, (ed.), *Rebirth of Rhetoric: Essays in Language, Culture & Education* (pp. 39-58). London: Routledge.
2. Terry Locke, (1995). The National English Curriculum of NZ: Has the Debate Ended, *English in Aotearoa*, 26, 44-49. Also, Terry Locke (1996) 'Yet again: Why the levels won't do, *English in Aotearoa*, 30, 28-32.
3. Moreover, as Bill Lennox notes in '*Advocacy, Evolution and Learning: School Assessment for the National Qualifications Framework*, a paper presented at the conference of NZARE, Palmerston North, 8 December, 1995, 'standards-based approaches' have been implicit in Bursary marking schedules for years.
4. Ministry of Education (1993). *Assessment: Policy to Practice*, Wellington: Learning Media, p. 10.
5. Department of Education, (1986). *Learning and Achieving: Second Report of the Committee of Inquiry into Curriculum, Assessment and Qualifications in Forms 5 to 7*. Wellington: Learning Media.
6. Prof. G. R. Hawke, (1988). *Report on the Working Group on Post-compulsory Education and Training*. Wellington: Learning Media.
7. See Warwick Elley, (1996). Curriculum Reform: Forwards Or Backwards?, *Delta: Policy and Practice in Education*, 48, (1) 11-18.
8. Bill Lennox, (1995). *Advocacy, Evolution and Learning: School Assessment for the National Qualifications Framework*, a paper presented at the conference of NZARE, Palmerston North, 8 December.
9. Lennox, p. 1.
10. Lennox, p. 21.
11. Lennox, p. 18.
12. Lennox, p. 20.
13. New Zealand Qualifications Authority, (1996). *How to Analyse Unit Standards: A Guide for English*. Wellington: NZQA. p. 3.
14. *How to Analyse Unit Standards*, p. 9.
15. New Zealand Qualifications Authority, (1997). 1997: *English Trials, Newsletter* 13. Wellington: NZQA. pp. 11-12.

16. Warwick Elley, (1996). Whatever happened to quality? *English in Aotearoa*, 30, 24.
  17. Elley, p. 23.
  18. Michael Peters & James Marshall, (1996). The Politics of Curriculum: Busnocratic Rationality And Enterprise Culture. *DELTA: Policy and Practice in Education*, 48, 1, (1996), 34.
  19. Elley, p. 24.
  20. Lennox, p. 21.
  21. John A. Codd, (1996). Higher Education and the Qualifications Framework, *DELTA: Policy and Practice in Education*, 48, (1), 60.
  22. Terry Hayes, (1996). Senior English in the Victorian Certificate of Education: a modern horror story. *English in Aotearoa*, 30, 12. Hayes's article is reprinted in *Idiom* (Journal of the Victorian Association for the Teaching of English), 32 (1), April 1997, which is a VCE Review Issue.  
Check<<http://www.srl.rmit.edu.au/vate/idiom/idiom2.htm>>
  23. *English study design*, Victorian Curriculum and Assessment Board, September, 1990, p. 42.
  24. *English study design*, p. 2.
  25. *English study design*, p. 3.
  26. The Hayes article provides a disturbing account of subsequent changes that have been made to the original study design as described in this article, especially in the CATs.
  27. Warwick Elley, Curriculum Reform: Forwards or Backwards?, pp. 15-16.
  28. Ministry of Education, (1997). *A Future Qualifications Policy for New Zealand: A Plan for the National Qualifications Framework, Green Paper*, Wellington: Learning Media.
-