REPORTING OF ACHIEVEMENT in
A COMPETENCY-BASED EDUCATION
AND TRAINING SYSTEM
with particular reference to
THE ISSUE OF "GRADED ASSESSMENT"
REPORTING OF ACHIEVEMENT

in

A COMPETENCY-BASED EDUCATION AND TRAINING SYSTEM

with particular reference to

THE ISSUE OF "GRADED ASSESSMENT"
Prepared for the

ASSESSMENT CENTRE FOR VOCATIONAL EDUCATION

by

ROBERT QUIRK

July, 1995

This paper was commissioned to explore issues involved in the assessment and reporting of achievement in a competency-based education and training system.

It is a contribution to informed discussion and debate.
ACKNOWLEDGMENTS

The reference group for the preparation of this report was the Assessment Policy Review Group of the NSW TAFE Commission, which consisted of:

- Cathy Barry (Chair), Curriculum Strategy Unit
- Mike Bailey, University of Sydney
- Lynda Henderson, Basic Work Skills Training Division
- Margaret Summerfield, Business Services Training Division
- Alison Bray, Accreditation Unit (Course Information Systems)
- Brian Thomas, Engineering Services Training Division
- Elizabeth Thilo, Pre-Vocational Programs Training Division
- Lindsay Cotton, Assessment Centre for Vocational Education
- Ron Wright, Southern Sydney Institute of TAFE
- Ric Moore, Northern Sydney Institute of TAFE
- Beryl Fitzharris, Sydney Institute of Technology
- Ingrid Zoebe, Accreditation Unit
- Kathy Wood, Student Administrative Services
- Katherine Rogers, Curriculum Strategy Unit
- Lee Ushakoff (Executive Officer), Curriculum Strategy Unit

The valuable work of Gary McIndoe, Research Assistant for the survey and interview components of the study, is gratefully acknowledged.

The author is grateful to all the informants to the project for their cooperation, to John McCrum, Assistant Director, Hunter Institute of Technology for his assistance in administering a survey to all teaching sections in the Institute's colleges, and to the staff of the TAFE Executive Library Service, Evan Bailey and Margaret Swinbourne, for their valued assistance.
# TABLE OF CONTENTS

<table>
<thead>
<tr>
<th>Section</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. Executive Summary</td>
<td>1</td>
</tr>
<tr>
<td>2. Recommendations</td>
<td>6</td>
</tr>
<tr>
<td>3. Purpose of the Study</td>
<td>8</td>
</tr>
<tr>
<td>4. Definition of Terms</td>
<td>10</td>
</tr>
<tr>
<td>5. The Issue of “Graded Assessment” in a System of Competency-based</td>
<td></td>
</tr>
<tr>
<td>Education and Training</td>
<td>11</td>
</tr>
<tr>
<td>5.1 Origins of the view that competency-based assessment should</td>
<td>11</td>
</tr>
<tr>
<td>recognise only a single level of achievement</td>
<td></td>
</tr>
<tr>
<td>5.2 The question of the compatibility of “grading” with the</td>
<td>13</td>
</tr>
<tr>
<td>principles that are fundamental to a competency-based education and</td>
<td></td>
</tr>
<tr>
<td>training curriculum</td>
<td></td>
</tr>
<tr>
<td>5.3 The question of whether or not, as a matter of policy, the option</td>
<td>13</td>
</tr>
<tr>
<td>to “grade” should be made available</td>
<td></td>
</tr>
<tr>
<td>5.4 The question of whether or not to “grade” in a given case, if a</td>
<td>21</td>
</tr>
<tr>
<td>“grading” option is available</td>
<td></td>
</tr>
<tr>
<td>5.5 The question of whether “graded” assessment should be based on</td>
<td>24</td>
</tr>
<tr>
<td>course/module learning outcomes or on other achievements</td>
<td></td>
</tr>
<tr>
<td>5.6 Methods of “grading” assessment in a competency-based education</td>
<td>25</td>
</tr>
<tr>
<td>and training context</td>
<td></td>
</tr>
<tr>
<td>5.7 The use of criterion-referenced assessment both to determine</td>
<td>30</td>
</tr>
<tr>
<td>whether or not an individual has achieved the outcomes specified as</td>
<td></td>
</tr>
<tr>
<td>essential for completion of a module or course, and to determine the</td>
<td></td>
</tr>
<tr>
<td>degree of merit, or grade, the individual has achieved</td>
<td></td>
</tr>
<tr>
<td>5.8 The question of obtaining both criterion-referenced information</td>
<td>33</td>
</tr>
<tr>
<td>(to recognise the achievement of essential outcomes) and norm-</td>
<td></td>
</tr>
<tr>
<td>referenced (for purposes such as ranking)</td>
<td></td>
</tr>
<tr>
<td>5.9 The form in which the results of “graded” assessment may be</td>
<td>38</td>
</tr>
<tr>
<td>reported</td>
<td></td>
</tr>
<tr>
<td>6. Method for the Study</td>
<td>40</td>
</tr>
</tbody>
</table>
7. Results of the Study .................................................. 42

7.1 Assessment of levels of achievement in a competency-based system of education and training ........................................ 42
7.2 Preferred method of reporting levels of achievement in final result notices for modules ................................................. 44
7.3 The use of grades for reporting on the assessment of levels of achievement ................................................................. 45
7.4 The use of marks for reporting on the assessment of levels of achievement ................................................................. 48
7.5 The use of verbal descriptions for reporting on the assessment of levels of achievement ................................................. 51
7.6 Reporting levels of achievement in a course ................................................................. 51
7.7 Perceptions of teachers regarding student responses to the recognition of merit ................................................................. 54
7.8 A note on the use of the terminology, “Pass/Fail” ................................................................. 55

8. Implications .................................................. 56

9. References .................................................. 59

10. Informants .................................................. 62

11. Appendix .................................................. 69
I. EXECUTIVE SUMMARY

An examination of relevant literature indicated that the historical origins of the practice of reporting only on a "pass/fail" basis (or its equivalent) in competency-based education and training are to be found in the combined influence of the "mastery learning" movement and criterion-referenced testing on competency-based education and training in the United States, and in the emphasis on access and equity in the subsequent development of competence-based National Vocational Qualifications (NVQs) in the United Kingdom.

The question arises as to whether it is necessary or appropriate for all curricula delivered by the vocational education and training sector under Australia's National Training Reform Agenda to be assessed only on a "pass/fail" basis (or its equivalent). The literature indicates:

- that a distinction should be drawn between the principles that are integral to a competency-based education and training system and other educational or philosophical principles.
- that the principles integral to competency-based education and training do not necessarily preclude "graded assessment";
- that the question of grading is, in fact, a policy decision.

In particular, two decisions about policy emerge.

The first is whether or not, as a matter of policy, to provide the option to "grade" in appropriate circumstances. Opposition to the provision of a "grading" option tends to be associated with arguments based on:

- a belief that "grading" is fundamentally incompatible with the principles integral to competency-based education and training;
- pragmatic considerations, such as a belief that, although admissible in principle, the provision of "grading" would tend to develop a "norm-referenced mindset" and undermine the criterion-referenced judgment that is integral to competency-based education and training, about whether or not the learning outcomes specified as essential for "passing" have been achieved;
- a belief that the educationally undesirable effects of grading are so common and so significant that assessment policy should prohibit "grading" in all cases;
- a belief that "grading" is not compatible with the operation of the vocational education and training system that has been established in Australia through the National Training Reform Agenda, including features such as endorsed competency standards, an Australian Standards Framework and an Australian Qualifications Framework.

Opposition on these grounds to a "graded assessment" option being available in the case of any
curriculum delivered by the vocational education and training sector under Australia's National Training Reform Agenda is not, in the main, supported in the literature. Whether the learning outcomes of a competency-based curriculum involve the achievement of "workplace competence", "workplace readiness" or a different kind of outcome, the principles of competency-based education and training require the assessment report to make clear whether or not a learner has met the essential requirements for the achievement of the learning outcomes. The literature suggests, however, that in many cases it would also be desirable to design assessment procedures to report on the learner's achievement of some degree of merit in a module or course. That would be compatible with the principles of competency-based education and training, and would essentially be a policy decision informed by educational considerations, client expectations, technical assessment considerations and other factors.

The second policy-related decision to be made, if the option to "grade" is available, is whether or not to "grade" in a given case. Such a policy decision would reflect the values and priorities of the policy makers and/or their clients:

- a decision not to "grade" could be based, for example, on the desire to develop cooperative rather than competitive attitudes through the course, or because, for the particular learners involved, "grading" may not provide significant benefits and may even be alienating; even in cases where learners will be facing a selection process, a decision not to "grade" might be taken based on a view that selection is an issue for the selectors (such as employers or educational institutions) and should not be allowed to influence the assessment processes for a course;

- a decision to "grade" could be based, for example, on grounds of access and equity, namely that if "pass/fail" assessments were imposed on qualifications in a system where the qualifications were being used for selection or progression, selection would be carried out much less fairly, based on factors such as the reputation of the college attended or the locations where particular individuals learned their skills.

The literature suggests that there will be cases where "grading" would have detrimental effects or be inappropriate for educational reasons. It suggests that where "grading" is an assessment option, available for use in cases where it serves a valuable purpose, assessment policy should ensure that it is used only in appropriate cases.

To implement any policy decision to "grade" assessment in a competency-based curriculum, however, the following issues need to be clarified:

- the basis on which criterion-referenced assessment would operate for the purposes of producing "graded" assessment reports, and in particular the extent to which "graded" assessment would need to be based on predetermined, detailed, documented, special learning outcomes and related assessment criteria;

- the extent, if any, to which it would be appropriate to supplement criterion-referenced assessment information in a competency-based education and training program with norm-referenced assessment information.
For the most part, the literature on "grading" is concerned with recognition of merit based not on extraneous criteria, such as how quickly an individual learns, but on the achievement of the learning outcomes specified for a module or course, through the assessment of achievement at levels beyond the standard set for a "pass".

It is emphasised in the literature that in a competency-based education and training program, judgments about whether or not an individual has achieved the outcomes specified as essential for completion of a module or course must be based on criterion-referenced information.

If a policy decision is made also to "grade" assessment in a particular course, the issue arises about whether or not the award of one or more degrees of merit (beyond the essential criterion-referenced judgement of "pass/fail" or its equivalent) should also be based on criterion-referenced information. The way this issue is resolved may vary depending on the purpose of "grading". For example, if the purpose is simply to encourage and reward excellence, a single merit grade may suffice, and this may be awarded on the basis of criterion-referenced information. On the other hand, if the purpose is to provide a basis for ranking individuals for purposes such as selection, multiple "grades" may be necessary and basing them on criterion-referenced information may be more complex and problematic.

Three broadly different approaches to this issue may be identified in the literature:

1. the use of criterion-referenced assessment both to determine whether or not an individual has achieved the outcomes specified as essential for completion of a module or course, and to determine the degree of merit, or grade, the individual has achieved;

2. the use of an assessment method that will provide both criterion-referenced information (to recognise the achievement of essential outcomes) and norm referenced information (for purposes such as ranking). This could involve:
   - the use of competency maps or competency continua, which may be based on the application of latent trait theory; or
   - the legitimate derivation of norm-referenced interpretations from assessments that are primarily criterion-referenced;

3. the use of criterion-referenced assessment (to recognise the achievement of essential outcomes) and separate norm-referenced assessment (for purposes such as ranking).

The project report has identified the exploration of the applicability and implementation of these three approaches as a matter for further research. Such research would also consider whether the sharp dichotomy that has tended to be drawn between norm-referenced and criterion-referenced assessment, in much of the discussion of "graded assessment" in a competency-based education and training system, is justified.

The report has also identified as matters for further research:

- an investigation of current assessment practice, with particular reference to the way it
might be informed and improved by an exploration of the above approaches to the "grading" issue;

- the technical assessment issues involved in a number of specific "grading" procedures, to advise practitioners on the most appropriate assessment methods to use for "grading" modules and for "grading" courses.

The research study carried out in the course of this project aimed to identify stakeholders' preferences regarding the methods of reporting on levels of achievement. It produced the following findings based on responses to questionnaires, interview and focus group questions by respondents including TAFE Training Division and teaching staff, and staff of private providers, government agencies, Industry Training Advisory Bodies and industrial enterprises:

1. The most widely supported view was that levels of achievement may be assessed in a competency-based system of education.

   The implementation of this principle would require one or more methods of reporting on levels of achievement to be available as an option for use in appropriate circumstances.

2. The most widely supported form for the reporting of levels of achievement on modules was grades.

   There was widespread support for supplementing the reporting of grades with reports of verbal descriptions of achievement and/or marks.

3. The survey found more widespread support for reporting levels of achievement through a fixed number of grades than for permitting the number of grades reported to be variable from module to module or from course to course.

4. The survey did not find a clear preference for a specific number of grades to be used at module level, except in the case of Training Divisions, where there was a preference for three grades.

5. There was no clear preference for a specific kind of grading terminology, except where only two grades were reported.

   In the case of two grades being reported, Training Division respondents had a clear preference for "Pass, Pass with Merit". Many Training Division staff, however, were concerned that confusion was likely to occur if the term "Pass" was used in reporting assessment results for ungraded modules, as well as to refer to the lowest grade in graded modules.

6. The survey did not find decisive support for or against the view that the use of marks was appropriate in competency-based education and training, except in the case of TAFE teachers: there was a clear preference amongst TAFE teacher respondents for the view that marks may be used in competency-based education and training.
7. There was a clear preference for the use of a range of 0-100 marks, if marks were used.

8. The surveys of industry, of TAFE teachers and of private providers did not find decisive support for fixing the mark representing "Pass" or its equivalent at 50%, 70%, 80% or 100%.

These surveys did not find decisive support for permitting the "Pass" mark to vary from module to module.

(This particular issue was not raised in the earlier survey of Training Divisions.)

9. If verbal descriptions were used to report levels of achievement in modules, only Training Divisions and TAFE teachers demonstrated a clear preference for any one kind of verbal description: Training Divisions and TAFE teachers preferred comments selected by the teacher(s) from a list of standard statements.

10. There was a clear preference that it should be possible for overall levels of achievement to be reported in awards for some courses.

11. There was a clear preference overall for the recognition of three levels of achievement at whole course level, although Training Divisions were more evenly divided between support for two levels and three levels.

12. There was a clear preference for the use of award grades as the method of reporting levels of achievement in whole courses.

13. If two award grades were reported, the preferred nomenclature was "Pass, Pass with Merit".

If three award grades were reported, the preferred nomenclature was "Pass, Credit, Distinction".

14. There was a clear preference not to report grade point averages, whether they were reported instead of award grades or as well as award grades.

15. The great majority of TAFE teachers in the Institute surveyed believe that reporting different levels of achievement in a course/program, instead of pass/fail, could make the course/program more attractive to prospective students. This was either the dominant or the unanimous view of respondents in all fields of study.

16. The great majority of TAFE teachers in the Institute surveyed believe that reporting different levels of achievement in a course/program, instead of pass/fail, could benefit the teaching/learning process. This was either the dominant or the unanimous view of respondents in all fields of study.
2. **RECOMMENDATIONS**

1. That the findings (see Executive Summary) of this project's survey on the preferences of stakeholders inform the development of the NSW TAFE Commission's policy on the reporting of achievement in competency-based modules and courses.

2. That research be undertaken to explore the applicability and feasibility of each of the following three possible approaches to "graded assessment":

   - the use of criterion-referenced assessment both to determine whether or not an individual has achieved the outcomes specified as essential for completion of a module or course, and to determine the degree of merit, or grade, the individual has achieved;

   - the use of an assessment method that will provide both criterion-referenced information (to recognise the achievement of essential outcomes) and norm referenced information (for purposes such as ranking). This could involve:
     - the use of competency maps or competency continua, which may be based on the application of latent trait theory; or
     - the legitimate derivation of norm-referenced interpretations from assessments that are primarily criterion-referenced;

   - the use of criterion-referenced assessment (to recognise the achievement of essential outcomes) and separate norm-referenced assessment (for purposes such as ranking).

   Such research should include consideration of whether the sharp dichotomy that has tended to be drawn between norm-referenced and criterion-referenced assessment, in much of the discussion of "graded assessment" in a competency-based education and training system, is justified.

3. That research be undertaken to explore current assessment practice, to determine the extent to which practitioners understand the principles and procedures associated with both criterion-referenced and norm-referenced assessment, and to consider ways in which current assessment practice might be informed and improved by an exploration of the three assessment approaches identified above.
4. That research be undertaken into the technical assessment issues involved in options such as the following, identified in NSW TAFE Commission (1994), to advise practitioners on the most appropriate assessment methods to use for "grading" modules and courses:

Options for grading modules:

- inference based on professional judgment with reference to module purpose;
- standard procedures for combining graded results within module;
- totalling and/or averaging marks.

Options for grading courses:

- results from all modules may be combined to determine overall result;
- results from selected modules may be combined to determine overall result;
- averaged marks of modules contributing to award;
- standard procedures for combining graded module results within course;
- major piece of work assessing integrated outcomes of course.
3. **PURPOSE OF THE STUDY**

This study was proposed by the Assessment Policy Review Group convened by the NSW TAFE Commission's Educational Development Division.

The background to the study proposal may be outlined as follows:

The introduction of Competency Based Training (CBT) in accredited courses had raised a number of issues in relation to assessment. Two critical areas were

- the recognition of merit in assessment
- the method of reporting on the recognition of merit

Assessment could result in a report that the learning outcomes specified for a course or program have or have not been achieved, but it might be possible to recognise the achievement of one or more levels of merit, or to provide some other form of additional information about the learner's achievements. The NSW TAFE Commission, partly as a result of strong representation from a number of industry groups, has permitted the recognition of merit in its assessment and reporting arrangements for CBT curriculum. The recognition of merit is not a requirement, merely an option that is available for assessment and reporting for use in appropriate circumstances.

Some debate had occurred about the number of levels of merit that should be distinguished and the form in which the achievement of merit would be reported. Concern had also been expressed that assessment information reported to students and others was not always meaningful, and there had been some consideration of the use of explanatory notes aligned to the achievement of module purpose.

The Assessment Policy Review Group which had been considering these issues suggested that a survey of employers and other users of TAFE NSW assessment reports should be conducted to inform the Educational Development Division when determining policies and procedures.

With reference to these concerns, the aim of the study was defined as being:

1. "to investigate what industry and others actually want from the reporting of assessment";

2. "to identify and evaluate the options for reporting achievement".
In investigating what information should be obtained from assessment and reported on, the study was to consider whether this should include:

- whether or not a student has achieved learning outcomes that meet the aims of a course or the purpose of a module?
- how well a student has performed in relation to those outcomes?
- evidence of achievement or attributes (knowledge, skills and/or attitudes) not necessarily limited to the learning outcomes pre-specified in course documentation?

In investigating the form(s) in which assessment outcomes should be reported, the study was to include the following options:

- if reporting on whether or not a student has achieved learning outcomes that meet the aims of a course or the purpose of a module:
  - Pass/Fail?
  - other?
- if reporting on how well a student has performed in relation to learning outcomes that meet the aims of a course or the purpose of a module:
  - grades?
  - marks?
  - verbal descriptions/profiles?
  - other?
- if reporting on evidence of achievement or attributes (knowledge, skills and/or attitudes) not necessarily limited to the learning outcomes pre-specified in course documentation:
  - verbal descriptions/profiles?
  - other?

In exploring these issues, the study would recognise that assessment and reporting could be confined to determining whether or not a student has achieved learning outcomes specified in syllabus documentation as meeting the aims of a course or the purpose of a module.

In exploring the two broad issues identified above, the study would identify any assessment and reporting requirements of particular stakeholders that might go beyond this minimum requirement.
4. **DEFINITION OF TERMS**

- A TAFE curriculum consists of one or more modules.
- A number of modules may constitute a whole course.
- Specified learning outcomes contribute to the achievement of a module purpose.
- Learning outcomes, module purposes and course aims are specified in a syllabus document.
- Assessment in a competency-based system is concerned with determining at least whether or not a learner has achieved essential curriculum outcomes (such as achievement of specified learning outcomes).
- There is some debate in the vocational education and training sector about whether it should also be concerned with how well a learner performs in relation to such curriculum outcomes; that is, about whether in some circumstances it should recognise different degrees of merit or levels of achievement.
- The survey is concerned only with reporting on curriculum outcomes in official NSW TAFE result notices, and does not address issues associated with other kinds of reporting, such as reports on attendance and cooperation that might be sent to employers of apprentices or of other students on training contracts.
- When discussing the form in which the results of assessment are reported, this paper prefers the term "grade" to refer to forms such as "A", "B", "C", or "Distinction", "Credit", "Pass", or "Merit", "Pass", which may be used to report the achievement of a small number (typically in the range of two to five) of different degrees of merit or levels of achievement attained by learners who have also attained the essential outcomes specified for a competency-based module or course.
- In the literature discussed in this paper, however, the terms "graded assessment" and "grading" occur frequently, and their use is not confined to assessment for the award of "grades" in the above sense. The use of the terms "graded assessment" and "grading" in the literature may encompass any form of reporting levels of achievement above the achievement of essential module or course outcomes, including the reporting of percentage marks. The term "grading" in particular is used as a convenient form of "short-hand" to refer to a wide range of merit-recognition practices. In discussing the relevant literature, the paper will necessarily at times employ the terms "graded assessment" and "grading" with this broad meaning.
5. THE ISSUE OF "GRADED ASSESSMENT" IN A SYSTEM OF COMPETENCY-BASED EDUCATION AND TRAINING

5.1 Origins of the view that competency-based assessment should recognise only a single level of achievement

The National Training Reform Agenda in Australia has been strongly influenced by work done in the United Kingdom associated with the establishment of National Vocational Qualifications (NVQs). As summarised by Shirley Fletcher (1992), the "competence-based" assessment approach associated with NVQs focuses on 'outcomes' rather than on the learning process. Moreover, competence-based assessment can be and should be undertaken in the workplace. It focuses on what the individual can do rather than what the individual knows. The approach focuses on determining that an individual has, by whatever means (which may or may not include the completion of a course), developed the ability to perform to industry standards. The award of an NVQ attests that the individual can perform to the standards for the relevant "competences".

Competence-based assessment for NVQs aims to report only that an individual is 'competent' or 'not yet competent': there are no percentage ratings; there are no comparisons with other individuals' results. (Fletcher, 1992)

Alison Wolf (1993) explores the historical origins of this preference for a "single cut-off" in assessment in competency-based education and training. She traces it to the close link that developed in the United States between criterion-referenced testing, as it grew in popularity, and 'mastery learning', an educational movement that aimed at improving the achievement level of a class as a whole. The 'mastery learning' movement in the United States had considerable success in improving the achievement of primary school classes in arithmetic and reading. It emphasised the achievement of a specified standard of performance in a subject by every member of a class.

In the United States, competency-based vocational education and training tended to be based on criterion-referenced assessment and 'mastery learning'. The rationale for this is outlined by Blank (1982): the standard for "mastery" is set at a high level for each task; this ensures that most students will attain a high level of proficiency on the task; this in turn will ensure a sound foundation for the learning and mastery of subsequent tasks; because mastery of each successive task is required for a student to progress through a training program, the competence of those who complete the program is assured; less able students may take longer to complete the program and require more help but will be as competent as "more able students" on completion.

In contrast, according to Blank, the effect of reporting grades or marks on how well a task has been performed is to encourage incompetence: for example if a percentage scale is used, and the minimum level for passing is set at 70%, this suggests that 30% of the program content is not very important; moreover this minimum level for passing will for some students become the maximum level of their performance.
Wolf (1993) expresses the opposing view, that is, that the standards set in "mastery" programs tend be not only minimum, but also maximum, levels of achievement:

"teachers teach to them, but no further, which depresses the achievement of the more able or harder working students."

According to Blank (1982), achievement of a standard of performance beyond that specified for mastery of a task constitutes achievement on a more advanced task, not superior achievement on the same task.

It may be noted that Blank (1982) does not exclude grading absolutely. He supports self-paced mastery learning, but, unlike many supporters of competency-based training, he argues that the time a learner takes to complete a program is significant. He favours a grading system based on the number of tasks mastered by a learner in a specified period of time: excellence is indicated by speed of learning.

In exploring the origins of the NVQ approach to assessment, Wolf (1993) points out that government policy in the U.K. was concerned to break down barriers to entry to certain occupations, especially craft occupations, and to increase the numbers of people with vocational qualifications. A strong commitment to access and equity principles influenced the development of a new system for the recognition of vocational skills. The system was based on the recognition of "competences", contained in standards of competence developed by industry lead bodies. An individual had to demonstrate specified "competences" in order to be awarded a National Vocational Qualification (NVQ). It did not matter if the individual had developed particular "competences" through a training program or through experience in the workplace. All that should matter was whether or not the individual was competent.

Graded assessment was considered to be one of the trappings of formal education that the NVQ system was reacting against. For example, Jessup (1989) argued that vocational education and training in the U.K. had been dominated by "an educational model of assessment" concerned with "discrimination between individuals", and in the past only 40% of the population had participated in such programs. As well as reducing the relevance of VET to employment, this model had the effect of excluding from VET people not suited to an 'academic' environment. The new VET model was designed for the whole population.

While an approach to assessment and reporting that recognises only a single level of achievement (for example, "competent"/"not yet competent", "achieved"/"not yet achieved", or "pass"/"fail") may be preferable for some courses, the question is whether it should be required of all competency-based curricula delivered by the vocational education and training sector.
5.2 The question of the compatibility of "grading" with the principles that are fundamental to a competency-based education and training curriculum

A distinction may be drawn between the principles that are integral to a competency-based education and training system and other principles to which particular individuals or groups using the system are committed and which they desire to see incorporated in the system.

Hager, Athanasou and Gonczi (1994) identify one argument against "grading" as being that "grading encourages competition rather than cooperation in learning". The educational process may be viewed as a vehicle for encouraging cooperative rather than competitive attitudes and values in the workplace and other social contexts outside the educational program. If competency-based programs report against targeted outcomes only in terms such as "achieved" or "not achieved", they may lend themselves to the promotion of cooperative values.

It is argued that these perceived advantages are associated with any non-graded assessment, whether it is competency-based or not. Competency-based assessment may be regarded as a form of criterion-referenced assessment, and, according to Wolf (1993), "grading and criterion-referencing are frequently discussed as though the two were antithetical". She demonstrates, however, that criterion-referenced assessment does not necessarily exclude graded assessment reports. (See section 5.6 below)

The principles integral to competency-based education and training do not necessarily preclude "graded assessment". As Hager et al. (1994) put it, "the question of grading is, in fact, a policy decision".

In particular, two decisions about policy emerge:

- the first is whether or not, as a matter of policy, to provide the option to "grade" in appropriate circumstances;
- the second, if the option to "grade" is available, is whether or not to "grade" in a given case.

5.3 The question of whether or not, as a matter of policy, the option to "grade" should be made available

When competency-based assessment was introduced in the U.K., it was argued, regarding the new system as a whole:

"New forms of assessment may take some time to establish their credibility, simply because they are new. People prefer to stick with methods that they know. They tend to be more comfortable with assessment methods they have undertaken themselves and use them as benchmarks of quality... Time and effort will be needed to promote the credibility of any new system, but that credibility will come with an
understanding and acceptance of the system's aims and how it works in practice." (Training Agency, 1988)

Similarly, it was suggested in the working group on excellence and expertise of the Assessment Steering Group, CBT Working Party (1993a) that moves to introduce a graded competency-based assessment system could be premature when a criterion-based assessment system has not yet been implemented. It could introduce "a further complication in the midst of a difficult change process already under way". More specifically, "many people are genuinely concerned that a graded assessment will reintroduce a normative assessment mindset and that important features of the competency-based training movement will be watered down or lost." It was believed that it would take time for people to get used to, understand and accept a criterion-referenced approach to assessment.

Such arguments allow that "grading" is compatible with the principles integral to competency-based assessment, provided they are derived from criterion-referenced assessment, but "grading" is opposed on the pragmatic grounds that it will too readily result in norm-referencing rather than criterion-referencing, and undermine the foundations of competency-based education and training. The underlying assumption is that norm-referenced assessment has no place within the vocational education and training sector under the National Training Reform Agenda.

A decision not to "grade" assessment could also be based on the grounds that it would be too expensive, or too cumbersome to implement. (Basic Work Skills Training Division, 1994; Hager et al., 1994) The force of this objection may depend on the methods used to design, implement, record and report "graded" assessment.

Apart from these practical considerations, however, it is argued that the educationally undesirable effects of grading are so common and so significant that assessment policy should prohibit grading in all cases. The following arguments identified by Hager et al. (1994) should be placed in this category:

- that grading "encourages competition rather than cooperation in learning";
- that grading "will inevitably lead to a feeling of failure amongst those that receive low grades. This would adversely affect learning."

There is also a view that "grading" leads to a concentration of teaching effort on the more able students. (Assessment Steering Group, CBT Working Party, 1993a)

As Hager et al. (1994) note, however, such objections to "grading" in some ways could be applied to any assessment system: they are not confined to competency-based assessment.

Some arguments for the universal prohibition of "grading" are, however, specific to competency-based assessment within the National Training Reform Agenda.
While "grading" may be compatible in principle with a competency-based approach to education and training, there is a question about whether or not "grading" is reconcilable with the vocational education and training system that has been established through the National Training Reform Agenda.

According to one view, the nature of the qualification awarded in the current national competency-based system of vocational education and training is fundamentally different from any previous kinds of qualifications awarded, and that the nature of the new qualifications is incompatible with the use of "graded" assessment.

This view appears to admit of no circumstances in which "grading" would be legitimate under the National Training Reform Agenda. It is a view that depends on a particular understanding of the relationship between competency standards and curriculum.

An explicit and comprehensive development of this argument was not found in the literature, and so the following summary of the argument leading to this view has been included here, derived largely from responses requested from informants to this study who opposed the adoption of "graded" assessment on these grounds:

- under the Australian Qualifications Framework (AQF), qualifications issued in the vocational education and training sector are based on competency standards that have been developed by industry (and sometimes by enterprises) and endorsed by the National Training Board;

- the Australian Standards Framework (ASF) distinguishes eight levels of competency standards;

- at any one level, competency standards do not provide for recognition of different levels of achievement or degrees of merit;

- the essential purpose, or the only purpose, of competency-based assessment is to determine whether or not the individual has achieved the standards, that is, whether or not the individual is competent;

- vocational education and training curriculum is about competency standards: courses are not just written with reference to standards, they enable the individual to achieve the standards;

- a vocational qualification certifies that an individual has achieved the standards: for example, an individual is awarded a Certificate III if he or she successfully completes a competency-based course focusing on standards set at the corresponding ASF level 3;

- this competency-based vocational education and training system requires courses to be modularised: completion of, or recognition against, an appropriate combination of modules constitutes completion of a course and entitles the individual to the award of the relevant qualification;
"grading" would subvert this system:

- "grading" would involve the identification of more than the eight levels specified in the ASF, or would represent assessment against more than one of the eight ASF levels;
- "grading" would undermine the integrity of the module, because an advanced grade in one module is likely to represent achievement of part of the learning outcomes of a more advanced module;
- one advantage of a modularised curriculum is that it creates the possibility for people to move through the curriculum quickly, progressing to another module as soon as they satisfy the requirements for the preceding module: "grading" would discourage people from progressing in this way, by providing them with an incentive to devote more time to each module in order to maximise their grades.

The assumption that it is reasonable to expect accredited vocational education and training courses invariably to "deliver" competency standards, in the sense that they would certify that holders of qualifications were "workplace competent" as described in the relevant competency standards is challenged in the literature. According to Hager et al. (1994):

"Usually standards for workplace assessments are set at the level of competency of a well-trained person coming to the task without extensive previous experience (i.e., entry-level). In some cases, much higher standards will be set."

The amount of workplace experience that an individual will need, after completing a full-time off-the-job course, in order to become "workplace competent" will vary with the individual, the course he or she has completed, and the characteristics of the workplace itself. It will range from none to considerable.

In his survey of competency-based vocational education in Australia, Watson (1993) found that the role of TAFE in the competency-based training agenda needed to be clarified:

"Most of the TAFE personnel interviewed admitted uncertainty about whether it was their role to develop workplace competence or to develop learning outcomes en route to competence. If it is the latter, TAFE can retain its traditional training role. If it is the former, however, it is recognised that TAFE will have to get involved in the provision of genuine work-based learning, providing appropriate on-the-job training and learning opportunities replicating the drama, variety and range of the workplace. This possibility of course poses very significant resource and curriculum problems for TAFE."
Similarly, Sheldrake (1993, in Hall, 1994) identified as a key issue needing clarification:

"Are we assessing competence as demonstrated or as potential?"

Steenholdt (1994) was concerned at a suggestion in a VEETAC discussion paper that the award of a qualification would require demonstration of workplace competence:

"This immediately raises issues of accessibility and hence equity. If credentials will only be issued after completion of a work experience component of training the question has to be raised as to whether this will exclude students who are not employed or who do not have access to such work experience."

In December 1993 Commonwealth and State Ministers endorsed a National Qualifications Framework, which provided for six levels of qualifications for the vocational education and training sector: Certificates (I, II, III, IV), Diploma and Advanced Diploma. Descriptors were provided for each qualification level to establish common ground for qualifications across all sectors providing post-compulsory education and training.

In discussing these descriptors, the National Qualifications Framework (1993) reiterated that the vocational education and training system in Australia was to be a competency-based system of training, based on national competency standards endorsed by the National Training Board, and it stated that "where these standards exist they must be achieved in order for a qualification to be issued."

Because the performance required by competency standards is performance to the specified standard of workplace competence, the National Qualifications Framework (1993) might be interpreted as requiring full attainment of workplace competence as a condition for the award of a qualification.

The achievement of workplace competence does not necessarily depend on actual workplace experience - there are many examples of workplace simulation to aid learning - but Steenholdt (1994) has argued that "some competencies, because of their contextual nature, their complexity or perhaps the long time-span required, cannot be directly duplicated in a training situation". In reality, attainment of workplace competence will often be achieved only after substantial practice in a workplace. This practice may be part of a structured integration of "on the job" and college-based learning; or it may occur only after the completion of a college-based course or program. Economic conditions and industrial practices may make it difficult for students to gain access to integrated on- and off-the-job training, and where students have access only to off-the-job training, it will not always be possible for them to develop workplace competence as described in a set of competency standards.

Johnstone (1993), who is the manager of the ACTRAC National Laboratory Science Technician Curriculum Project, argues that a gap necessarily exists "between the
totality of any set of competency statements and what can be delivered and assessed off-the-job by any provider. In the case of individual learning outcomes, the skills of learners on entry to an exclusively off-the-job learning program, the time and resources available for the program and its essentially preparatory nature tend to result in a difference between the level of performance specified for attainment in the learning program and the level of performance specified in the competency standards. Attainment of workplace competence would require learners to have the opportunity subsequently to gain access to a workplace, normally through employment, to extend and practise their skills over time.

In the vocational education and training sector, while the recognition authorities regulate the issuing of qualifications, providers usually issue testamurs. It would normally be a goal of vocational education and training programs that their graduates will have achieved as much of the relevant competency standards as possible. According to Byrne (1994), if qualifications were issued only on achievement of full workplace competence, many off-the-job courses would not lead to a full qualification. This raises the problem of who would issue the qualifications where achievement of full competence was subsequently achieved when the individual was in employment. The issuing authority might have to be the relevant industry or professional body. If this approach were adopted, vocational education institutions such as TAFE could become merely providers of programs, completion of which would be accepted as evidence of partial progression towards the achievement of full workplace competence. (Byrne, 1994)

In contrast, a respondent to the study that is reported in this paper referred to certain professions, arguing that people learning law, medicine or dentistry could not be registered or licensed to practice until they had completed "a post-course on-the-job component". This argument was advanced in support of the view that the award of a vocational qualification should represent achievement of "workplace competence". In none of the professions cited, however, is possession of the educational qualification a license to practice. In all cases the qualification is issued by the university independently of the granting of the right to practice. In two of the three cases (law, medicine), admission to the profession is not granted until the applicant has successfully completed a period of professional development (College of Law, medical internship) that follows the award of the educational qualification, - that is, full occupational competence is gained after the award of the qualification: the qualification itself does not certify full occupational competence. (Tom Ruben, Roger Wescombe, and Hugh Wilson: personal communications) If anything, the comparison with the professions corroborates the need to distinguish between meeting the requirements for the award of a vocational qualification by an educational institution and the attainment of "workplace competence".

Insistence that the award of a vocational qualification by a TAFE authority should depend on the prior attainment of workplace competence would give rise to equity concerns similar to those identified by Steenholds (1994), given the scale and scope of TAFE's provision in post-compulsory education for people in employment not related to the vocational course or program in which they are enrolled, people in related
employment but without adequate workplace learning opportunities, and for many people in full-time education where workplace simulations may not be feasible. The TAFE system provides off-the-job training to both full-time and part-time students. A profile of TAFE’s clientele is indicated in Dawe (1993), where only 40% of respondents to a national survey of TAFE students appeared to be undertaking courses related to current employment. 26% were preparing for future employment.

In the case of NVQs in the U.K., it would have been inevitable, given the focus of NVQs on workplace competence, that a large number of students in further education would be similarly disadvantaged. The award of NVQs tends to depend on individuals having had substantial workplace learning and having been assessed in workplaces or "realistic workplace environments".

To address this issue, a second kind of qualification has now been introduced into the NVQ framework: the General National Vocational Qualification (GNVQ). The introduction of GNVQs represents a recognition of the need to provide for the full range of people wanting access to vocationally relevant qualifications. In July 1993, the Government formally identified NVQs as "the vocational route mainly for those who have left full-time education'...[and] GNVQs as the vocational route mainly for those in full-time education, and described advanced GNVQs as 'vocational A-levels'" (Smithers, 1993)

Like NVQs, GNVQs are based on standards written in outcome terms; unlike NVQs, however, they do not aim to achieve full workplace competence. Learning outcomes are in the form of "statements of achievement" rather than "statements of competence".

Moreover, in the case of GNVQs, provision is made for the reporting of "graded" assessment. (Business and Technology Education Council, 1993)

According to Wolf (1995), "the vast majority of NVQs awarded are at Levels 1 and 2. At technician and business management level, there has been considerable resistance to NVQs, with the older BTEC Diploma awards holding their ground. It seems likely that these diplomas will in fact be incorporated into the more traditional 'educational' GNVQ (General National Vocational Qualification) system rather than ever becoming fully fledged competence-based awards."

The fact that the National Qualifications System in Australia provides only one set of qualifications, in contrast to the dual system of NVQs and GNVQs operating in the NCVQ system, indicates that making the award of vocational qualifications conditional upon the attainment of full workplace competence in all cases would be unrealistic, and may suggest that a more flexible approach to the development of competency-based curriculum and the award of qualifications is inevitable. It is, however, essential for the assessment and reporting of curriculum outcomes to avoid confusion. Where achievement of specified curriculum outcomes also involves achievement of workplace competence as described in competency standards, this needs to be explicit. Where achievement of specified curriculum outcomes involves the achievement of knowledge and skills needed in the workplace but not yet developed to the level of workplace
competence (that is, what some curriculum developers refer to as "workplace readiness"), this also should be made clear.

With reference to the issue of "graded assessment", the question is whether provision should be made in the Australian competency-based education and training system for "graded assessment", paralleling the provision of "graded assessment" in GNVQs in the competency-based education and training system that has evolved in the U.K. under the National Council for Vocational Qualifications.

In some courses, college-based training in realistic workplace-simulated environments may result in trainees being "workplace competent" at the commencement of their employment. It is a desirable goal for vocational qualifications to "deliver" the relevant competency standards in this sense. In many courses, however, training will have provided learners with the knowledge and skills that make them "ready for the job", rather than "workplace competent". They are equipped to progress efficiently through practice and experience in the workplace to the achievement of "workplace competence" as described in the relevant competency standards. The question is whether it may be especially in these circumstances that a prospective employer would value a report on an applicant's achievement of merit in his or her course.

Recently the National Training Board publication, NTB Network, posed the question, "Is competency only assessed on a pass/fail basis?". The article stated:

"Grading of outcomes can occur against standards if there is a preparedness to write separate criteria for each grade. In reality, though, all this does is create more levels of achievement against which the judgement of competent or not yet competent can be made. Many organisations or educational institutions providing training or education against competency standards find that the pass/fail mark for a course, or the grades, can more effectively be determined against measures the organisation itself has added to the program, whether criterion based or normative. This could be based on achievement of additional outcomes, speed of learning compared to others in the group, and the mode of learning used. The point about them is that they relate to achievement in the learning environment, not necessarily on the job."

In contrast, however, to the reference in the NTB publication to "criterion based or normative" measures in education and training programs, a working group on excellence and expertise convened by the Assessment Steering Group, VEETAC CBT Working Party (1993b) stated:

"Any grading of the achievement of either learning outcomes or supplementary information in the learning modules, for learners whose competency has been established through an education provider, should be criterion referenced to be compatible with the philosophy of competency-based training."
This working group also commented:

"Industry recognised that it may be appropriate for providers to grade assessment for the recognition of excellence. However industry participants considered that it would be inappropriate to differentiate between workers on the basis of levels of competency".

Whether the learning outcomes of a competency-based curriculum involve the achievement of "workplace competence", "workplace readiness" or a different kind of outcome, the principles of competency-based education and training require the assessment report to make clear whether or not a learner has "passed" or "achieved" the learning outcomes, that is, whether or not the learner has met the essential requirements for the achievement of the learning outcomes.

The literature suggests that in many cases it would also be desirable to design assessment procedures to report on the learner's achievement of some degree of merit in a module or course. That, as has been seen, would be compatible with the principles of competency-based education and training, and would essentially be a policy decision informed by educational considerations, client expectations, technical assessment considerations and other factors.

5.4 The question of whether or not to "grade" in a given case, if a "grading" option is available.

In their discussion of "grading" Hager et al. (1994) comment that

"there has been an unfortunate tendency for the debate about grading to become polarised; people tend either to be vigorously in favour of grading or implacably opposed to it. This polarisation is unfortunate because it is possible to be both in favour of grading and opposed to it depending on the circumstances."

If a provider's general assessment policy makes "grading" available as an option, a decision about whether or not to "grade" assessments in a particular case would have to be justified by identifying the additional benefits that would result. According to Hager et al. (1994), this would require attention to factors such as:

- need to provide more comprehensive information;
- motivational impact;
- effect on learning;
- impact on validity and reliability estimates;
- feasibility and implications of its introduction;
- pressure from commerce and industry;
- higher education requirements for articulation, advanced standing or competitive entry.

In the case of particular courses or particular circumstances, it may be that grading has
educationally undesirable effects that outweigh any advantages of grading, and that assessment policy should ensure that grading does not occur in those cases.

Hager et al. (1994) give an example of a situation where the effects of grading could be seriously dysfunctional:

"In a TAFE Advanced Certificate for youth workers... where the learners were employed, were mature, had little formal education and were lacking in confidence, it could be argued that their capacity to progress would be adversely affected by grading, because it might lead to a sense of failure for those who performed badly. In this case written comments about how to improve would be more valuable than grades."

It may also be inappropriate to use graded assessment with certain kinds of content. For example, Basic Work Skills Training Division (1994) considers that the "emphasis on the teaching and learning of communication skills as an interactive and dynamic process means that the communicative competence of students may not be most appropriately or validly assessed by the application of degrees or grades of competence."

The literature suggests that the benefits of grading should not simply be assumed. For example, it may be widely assumed that graded assessment motivates learners to achieve their full potential, but Hager et al. (1994) state that "there is a good deal of evidence, from courses which do not have grading, which suggests that learners will perform to the maximum of their potential without grades." Similarly, it may be assumed that graded assessment reports are especially valuable to learners' current or prospective employers, but Basic Work Skills Training Division (1994) suggests that "the extent to which employers do actually make employment and promotion decisions based on TAFE grades is not entirely clear at present, and it will need to be pursued by further research."

There may be alternative methods of achieving some of the benefits intended to be achieved by "grading". The working group on excellence and expertise of the Assessment Steering Group, CBT Working Party (1993a), for example, reported "a view that individuals who demonstrate excellence are presently rewarded and recognised in many ways. In competency-based training they are rewarded by faster mastery of learning outcomes and can thus finish their training more quickly. In a competency-based training system, the trainer/assessor can provide feedback to the individual on how well they went. Such feedback is essential and can be given orally during the program. Furthermore they note that in life and work there are tangible rewards for excellence in terms of greater status, higher pay and faster promotion."

On the other hand, Woodrow (1994), who was investigating assessment in the automotive industry, suggests that employers who want grading believe that pass/fail reporting results in a reduced pressure to excel, and they want a better basis for gauging their apprentice's progress. At the same time, however, they want explanatory information on assessment gradings, as well as identification of those parts of a course
that their employees have had difficulties with. These employers use grades or marks for ranking apprentices, for selection for in-house and state apprentice awards, for planning additional training for poor performers, and for awarding cash bonuses.

Basic Work Skills Training Division (1994) believes there is a need to "review the impact of reporting performance under a competency based assessment scheme on articulation to higher education, particularly if a single pass result is reported", a need to "negotiate with these institutions for acceptance of a single pass grade for some modules as part of general articulation requirements."

While this could be negotiated for "some modules", it is evident that for those advanced TAFE courses from which some students do proceed to higher education, there is a significant selection problem that graded assessment helps to resolve.

Published information on detailed articulation arrangements between TAFE courses and university degrees in N.S.W. contains frequent references to grades both at whole course and at subject/module level. (NSW TAFE Commission, 1992) Dawes, Athanasou and Skelsey (1993) have investigated the selection of TAFE qualified applicants into university courses and found that, although TAFE grades may be used in conjunction with other criteria for selection, grades are highly valued in the selection process. These authors propose for consideration the use of a grade point average calculated on the basis of centrally examined TAFE subjects. They argue that grading "indicates past performance in a formal learning situation, usually in a similar field of study. Performance at one level of education is traditionally used as a selection criterion for another. Furthermore, affectively, it seems fair to reward those who perform well at one level of education with entry into a higher level."

According to the Principal Officer (Joint Secondary Schools/ TAFE; TAFE HSC Pathway) students applying to enter higher education based on an HSC program containing a TAFE vocational component would be seriously disadvantaged unless graded assessment was provided; indeed, "at this point in time and for the foreseeable future the ability to derive a mark for reporting to the Board of Studies is critical if vocational courses are to take their place in NSW as an integral component of HSC offerings which contribute to university entrance." (See Appendix 1)

The literature suggests there will be many cases where "grading" would have detrimental effects or be inappropriate for educational reasons. It suggests that where "grading" is an assessment option, available for use in cases where it serves a valuable purpose, assessment policy should ensure that it is used only in appropriate cases.

To implement any policy decision to "grade" assessment in a competency-based curriculum, however, the following issues need to be clarified:

- the basis on which criterion-referenced assessment would operate for the purposes of producing "graded" assessment reports, and in particular the extent to which "graded" assessment would need to be based on predetermined, detailed, documented, special learning outcomes and related assessment criteria

...
the extent, if any, to which it would be appropriate to supplement criterion-referenced assessment information in a competency-based education and training program with norm-referenced assessment information.

5.5 The question of whether "graded" assessment should be based on course/module learning outcomes or on other achievements?

Peddie (1993b) argues that there are two broad options for an assessment process to recognise the achievement of merit:

- recognition of merit based on the achievement of the learning outcomes specified for a module or course;
- recognition of merit based on factors beyond the achievement of the learning outcomes specified for a module or course.

Recognition of merit based on the achievement of the learning outcomes specified for a module or course could involve:

- achievement of learning outcomes required for a more advanced module or course;
- achievement at levels well beyond the standard set for a "pass";
- speed (attaining learning outcomes at a faster rate);
- consistency of performance.

Recognition of merit based on factors beyond the achievement of the learning outcomes specified for a module or course could involve:

- transfer of skills to new situations;
- achievement of additional learning outcomes;
- originality, creativity, 'flair';
- outstanding attitudes, approach to learning, motivation;
- merit awarded to the top 5 or 10 percent.

The achievement of additional learning outcomes has been considered by a number of writers. Additional learning outcomes are viewed as distinct from achievement on essential learning outcomes at levels well beyond the standard set for a "pass"; the distinction is not completely clear, however, as the additional learning outcomes would normally be related to the essential learning outcomes, and may represent more advanced applications of the essential learning outcomes.
Hager et al. (1994) are concerned that grading on extra work could be based merely on "diligence" rather than on the achievement of higher quality. Woodrow (1994) considers that grading on extra work would allow students to demonstrate more advanced skills and aptitudes, and could allow students to pursue their areas of interest, but also acknowledges that if, as was likely, project work were the basis for merit grades, assessment could be time-consuming, costly, subjective, and even invalid (given that, because project work was not tightly supervised, the extent to which a project was the outcome of a particular student's work could be difficult to determine); project work could also encroach on class time needed to teach essential outcomes.

A more acceptable option, in respect of "grading" achievement in a whole course, may be to include in the course structure a project to be graded which would enable the learner to demonstrate an integrated understanding and application of knowledge and skills learned across the course, and which would be assessed and graded holistically. (Byrne, 1993) Such a project could constitute either a core module or an elective module in the course structure.

The time taken to master specified learning outcomes has also been considered by a number of writers as a basis for grading. As has been noted, it is favoured by Blank (1982) on the grounds that slow learning is associated with unproductive behaviour, and that promoting productive behaviour by rewarding speed in learning with a grade promotes the competence of the trainee. To Hager et al. (1994), however, such an approach is both "at odds with an outcomes based system" which should not take the history of learning into account in assessment, and "too simplistic to be used widely as a measure of merit".

Woodrow (1994) identifies factors such as the following as possible bases for grading: use of safety procedures; efficient use of class time; punctuality; ability to follow instructions; ability to work with others; care of equipment/tools and college building; pride in workmanship; attendance. Some of these items are reminiscent of matters that have been reported on verbally in college reports to employers of apprentices. As factors to be included in the basis for grading, however, they have a number of disadvantages, most importantly that "many of these type of criteria would involve considerable subjective judging by the teaching staff".

For the most part, the literature on "grading" is concerned with recognition of merit based not on such extraneous criteria, but on the achievement of the learning outcomes specified for a module or course, through the assessment of achievement at levels beyond the standard set for a "pass".

5.6 Methods of "grading" assessment in a competency-based education and training context

Much of the discussion of assessment in competency-based education and training affirms that assessment should be "criterion-referenced" and not "norm-referenced". It is important to consider the relationship between competency-based assessment and criterion-referenced assessment, to examine the capacity of criterion-referenced
assessment to report on different levels of achievement or degrees of merit, and also to examine the assumption that norm-referenced assessment has no place within a broad system of competency based education and training.

Competency-based assessment is best understood as a particular type of criterion-referenced assessment. (Wolf, 1995) Criterion-referenced assessment "refers to the collection of evidence about a person's performance and judging the nature and extent of progress towards the requirements specified in clearly defined performance criteria or assessment criteria". (Hager et al., 1994)

Within the broad framework of a criterion-referenced approach to assessment, competency-based assessment focuses especially on the assessment of "real life", "occupational" performance as distinct from "academic" performance. (Wolf, 1995) It has emphasised assessment of performance in the workplace, typically by direct observation. Originally, the view was taken that if competent performance could be demonstrated, the knowledge contributing to the performance need not be assessed and could be inferred. Assessment of knowledge became acceptable in circumstances where it was not practicable to assess all possible variations in practice through performance demonstrations: supplementary assessment of the learner's underlying knowledge was needed to provide evidence that he or she would be capable of responding to a variety of performance demands. (Jessup, 1991)

The assessment of knowledge or theory components of a course might be thought to lend themselves especially to graded assessment. Is it the case, however, that demonstrations of performance, including performance in the workplace, do not lend themselves to graded assessment? Gonczi et al. (1995) provide an example of a Sydney Electricity assessment instrument for apprentices' performance that rates the "key job behaviours" using the following descriptions of levels of performance:

- "unacceptable performance: does not demonstrate necessary competency";
- "performance needs improvement: demonstrates potential and ability to develop competency";
- "satisfactory performance: competent, meets job requirements";
- "good performance: meets and at times exceeds job requirements";
- "very good performance: regularly exceeds job requirements";
- "exceptional performance: consistently exceeds job requirements, out of the ordinary".
A common misunderstanding of criterion-referenced assessment is that the term "criterion" refers to a standard (that is, a fixed level of performance) or a cut-off. In practice, criterion-referenced assessment has tended to adopt a "mastery" approach to testing, which in the case of competency-based assessment involves making a single judgment, such as whether the individual is "competent" or "not yet competent". As a result, there is a tendency to assume that "grading" is incompatible with criterion-referenced assessment:

"The majority of writers on the subject of competency based training and assessment ... suggest that competency based assessment should be criterion-referenced, and thus by implication should not return a graded result." (Basic Work Skills Training Division, 1994)

In principle, however, criterion-referenced assessment can recognise various levels of achievement, using a number of "cut-offs" (Wolf, 1995)

To some extent a 'mastery" approach to criterion-referenced assessment has been encouraged by a misunderstanding of the meaning of the term "criterion" in this context. Hambleton and Rogers (1991) identify Glaser (1963) and Popham and Husek (1969) as the seminal authors on criterion-referenced assessment, and they point out that these authors use the term "criterion" to refer only to "a domain of content or behaviour" to which test scores are referenced. A standard, in the sense of a fixed level of performance, need not be set on a criterion-referenced test. The following is perhaps the most widely accepted definition of this approach to assessment:

"A criterion-referenced test is used to ascertain an individual's status with respect to a well-defined behavioural domain" (Popham, 1978, in Hambleton and Rogers, 1991)

For example, a criterion-referenced test could be used with reference to the domain of content that addresses the knowledge and skills used in electron microscopy by a technician working in a medical laboratory. It could be reported that a particular learner performed satisfactorily on 75% of the test items. Provided that test development sampled appropriately from the domain, such a report would provide meaningful information about the individual's status with respect to that domain. This report would be meaningful, even though no "standard" has been set for performance on the test as a whole. (cf. Hambleton and Rogers, 1991)

"It is the ability of the test to give information about the examinee's status with respect to the behaviours in the domain that makes a test criterion-referenced, not the fact that the test is used with a cut-off score." (Nitko, 1984)

Criterion-referencing does not require a passing score to be specified for test scores to be meaningful, but a passing score may be specified if it is needed for certain purposes. (Nitko, 1984) Competency-based assessment does not require the setting of a passing score at 100%, as supporters of "mastery learning" might advocate. As Hager et al.
(1994) note, however, care is needed when setting a cut-off score; technical methods (e.g., Angoff, 1971) have been developed to determine the "cut-off" scores in criterion-referenced tests that correspond to the achievement of minimal competency or a "pass". As these authors also note, when a scale is selected to categorise performance, the use of nominal scales "which categorise performance and which classify it into pass-fail, competent/not-yet competent, or even into discrete grades such as pass or pass with merit, are consistent with competency-based and workplace assessment approaches".

Alison Wolf (1993) makes explicit the capacity of criterion-referenced assessment to report different levels of achievement or degrees of merit:

"Criterion-referenced assessment reports on performance in relation to an externally defined domain, but that does not mean that the only result is whether the performance matched the domain (achieved criterion level) or did not... Criterion-referenced assessments produce a 'distribution' of performance (or, in a formal sense, marks) in exactly the same way as any other assessment does. A single pass-fail line is ONE way to partition that distribution but only one."

She argues too that the number of cut-off points recognised should depend on the nature of the content of the program ("whether there are key, recognisable thresholds"). In fact, in the case of some outcomes such as "core skills" (or "key competencies"), assessment reports which recognise several grades of performance may be "much more true to the behaviour concerned".

Similarly, Wilmut and Macintosh (1994) argue that scope for differentiated outcomes from criterion-referenced assessment (such as grades) depends on what is being assessed. In their view, assessment of narrow task skills may be reportable only as "pass or not pass, or can do or cannot do", rather than in terms of degrees of merit or levels of achievement. They suggest that it is only when assessment is of more general attributes, and when it is possible to take a more continuous view of competence, that it will be possible to "determine attributes which characterise performance which goes beyond that required simply to pass, and which would be recognised as meriting greater recognition. In doing this we have moved beyond the elementary concept of mastery."

Peddie (1993a) demonstrates, however, that this continuous view of competence is not only applicable to more general attributes such as "key competencies". He illustrates a basis for recognising merit in typing: "If a Level one typing unit includes as an outcome the ability to type at twenty words per minute, and Level two is set at forty, then the learner who shows they can type at thirty w.p.m. may well be a candidate for a merit award at Level one."

The number of cut-off points recognised may also be influenced by the purposes for which the assessment reports may be used (Hager et al., 1994; Wolf, 1993), and by practical considerations such as feasibility, cost-effectiveness and the need for assessor training. (Hager et al, 1994)
Because of the strong policy preference for a single "cut-off" in the earlier development of competency based assessment in the U.S.A. and the U.K., there are very few precedents for alternative approaches. There is concern about the acceptability in principle of alternative approaches, especially if it is intended to distinguish enough levels of achievement to permit ranking for purposes such as selection. Peddie (1993b), for example, believes this may be "contradictory in terms of the competency-based approach". He goes on to suggest that "it may ... be desirable and even necessary for those groups with an interest in selection to promote alternative or subsequent assessments with the specific purpose of ranking learners".

From the previous discussion, however, it would seem that:

- the recognition of different levels of successful achievement in a course is compatible with competency-based education, irrespective of the effect of such recognition (for example, ranking) or its purpose (for example, selection), provided that a criterion-referenced judgment is made about whether or not each individual learner has achieved the outcomes specified as essential for the module or course;

- whether or not different levels of successful achievement are recognised is essentially a policy decision. Such a policy decision would reflect the values and priorities of the policy makers and/or their clients:
  - a decision not to "grade" could be based, for example, on the desire to develop cooperative rather than competitive attitudes through the course, or because, for the particular learners involved, "grading" may not provide significant benefits and may even be alienating; even in cases where learners will be facing a selection process, a decision not to "grade" could be taken based on a view that selection is an issue for the selectors (such as employers or educational institutions) and should not be allowed to influence the assessment processes for a course;
  - a decision to "grade" could be based, for example, on grounds of access and equity, namely that if "pass/fail" assessments were imposed on qualifications in a system where the qualifications were being used for selection or progression, selection would be carried out much less fairly, based on factors such as the reputation of the college attended or the locations where individuals learned their skills (Wolf, 1993).

It is emphasised in the literature that in a competency-based education and training program, judgments about whether or not an individual has achieved the outcomes specified as essential for completion of a module or course must be based on criterion-referenced information.

If a policy decision is made also to "grade" assessment in a particular course, the issue arises about whether or not the award of one or more degrees of merit (beyond the essential criterion-referenced judgement of "pass/fail" or its equivalent) should also be
based on criterion-referenced information. The way this issue is resolved may vary depending on the purpose of "grading". For example, if the purpose is simply to encourage and reward excellence, a single merit grade may suffice, and this may be awarded on the basis of criterion-referenced information. On the other hand, if the purpose is to provide a basis for ranking individuals for purposes such as selection, multiple "grades" may be necessary and basing them on criterion-referenced information may be more complex and problematic.

Three broadly different approaches to this issue may be identified in the literature:

- the use of criterion-referenced assessment both to determine whether or not an individual has achieved the outcomes specified as essential for completion of a module or course, and to determine the degree of merit, or grade, the individual has achieved;

- the use of an assessment method that will provide both criterion-referenced information (to recognise the achievement of essential outcomes) and norm referenced information (for purposes such as ranking);

- the use of criterion-referred assessment (to recognise the achievement of essential outcomes) and separate norm referenced assessment (for purposes such as ranking).

5.7 The use of criterion-referenced assessment both to determine whether or not an individual has achieved the outcomes specified as essential for completion of a module or course, and to determine the degree of merit, or grade, the individual has achieved.

If a number of degrees of merit, or grades, were to be recognised, and if outcomes and assessment criteria for each grade had to be specified and documented, the process might be time-consuming and costly, and could ultimately be futile.

Wolf (1993) demonstrates that the assessment of NVQs has been based on "continued commitment to criterion-referencing's promise of transparency, and a reliance on the standards alone to provide this", involving an "attempt to map out free-standing content and standards" that has led "again and again, to a never-ending spiral of specification". This has occurred merely in the attempt to provide a basis for making pass/fail judgements. If additional specifications had to be provided for each additional grade that was recognised, the problem would be compounded enormously.

The approach may be feasible if grading is based on a small number of simple variables. For example, the following option was developed for consideration in automotive programs:

- Each student who has achieved at least the specified essential program outcomes is awarded one of four grades which clearly defines the standard he or she has achieved within the unit. These four levels of achievement are:
Completes task on first attempt, meets all performance criteria, demonstrates high accuracy and efficiency.

Completes task on first attempt, meets all performance criteria, meets industry demands for efficiency and accuracy.

Completes task within 2 attempts, demonstrates accuracy and efficiency.

Completes task within 2 attempts, meets performance criteria within time requirements. (Woodrow, 1994)

Woodrow (1994) acknowledges that, despite the simplicity of this option, considerable time and energy would need to be given to the development of assessment criteria defining outcomes such as ‘high accuracy’, ‘within the timeframe required’ and ‘efficiency’ in relation to these four grades. If the recognition of degrees of merit were to be based on the assessment of a more complex concept of merit than this, even if fewer than four grades were to be recognised, the problem of specifying assessment criteria could be formidable.

Further, Wilmot and Macintosh (1994) argue that the reliability of assessment is likely to be compromised in a multi-grade model of assessment. They argue, for example, that "assessors cannot easily distinguish standards specified at several levels", that "assigning performances to categories is not a completely certain matter", and that "it is difficult to specify criteria which can be used for decision-making on specific pieces of evidence when assessment on several aspects of performance is incorporated into a single description".

These concerns, however, are based on the assumption that assessment of a number of different levels of achievement or degrees of merit would necessarily require the prior specification of detailed assessment criteria for each level of achievement that could be reported.

Hager et al. (1994) argue that the recognition of degrees of merit or grades, while "an impossible task in a 'clarity' model", would be feasible in a model incorporating a more integrated approach to competency, and specifying relatively few learning outcomes and making use of holistic assessment criteria. For example, Gonczi, Hager and Palmer (1994) report on the experience of a competency-based assessment exercise in the field of communication in the legal profession, based on the assessment of videotaped role play simulations by trainee assessors:

"In this training session, examples of the simulation videos were played and discussed by all assessors together and criteria for what constituted unsatisfactory, satisfactory and 'easily' satisfactory performance agreed on. All assessors had previously read the standards and were expected to be reasonably familiar with them. Agreement on the performance required for the satisfactory category was achieved very quickly.
suggesting that experienced practitioners have clear tacit models of professional performance. Consensus on the third category while not instantaneous, was achieved quite quickly suggesting that grading is not as difficult a task in criterion referenced assessments as many fear.

Similarly, Wolf (1993) does not believe that the introduction of more than one "cut-off" point requires the documentation of a set of detailed performance criteria for each "cut-off" point. She agrees that assessors do have difficulty assigning particular instances of performance to given categories of performance, but they find it much easier to rank performance:

"Having only a limited number of criteria or domain specifications on offer, does not stop people from ranking performance in terms of them and doing so with surprising consistency."

Where assessment is to be reported in grades, some specification of outcomes and assessment criteria for each grade may be appropriate, but it should avoid becoming lengthy, highly specific and constraining. According to Wolf (1993):

"people actually operate with complex holistic models of performance, in which they compensate and weight without being conscious of how and when they are doing so."

Clear specifications, Wolf argues, may be helpful but they are no substitute for this process. Valuable use may be made also of exemplars of achieved outcomes to help achieve consistency in assessment. Moreover, any system that aims to achieve clarity and consistency has to develop effective networks of assessors and teachers. (Wolf, 1993)

Networks of teachers and assessors can build a common understanding of the outcomes to be expected, both for pass/fail judgments and for multi-grade assessment. Such networks can discuss exemplars of outcomes and reach consensus on the nature of the key aspects that determine particular assessment judgments.

A multi-faceted approach to assessment is needed. Clear specification of outcomes is valuable, but it is

"only one aspect of sound and fair assessment. The use of exemplars and the building up of case law; the socialisation of assessors and their constant re-socialisation; the monitoring of marking reliabilities - all these are equally important." (Wolf, 1993)

If the purpose of reporting a number of grades of assessment is to rank individuals in order of merit for purposes such as recruitment or selection into further education, then this can be done by assessors in addition to reporting that particular individuals have achieved the essential outcomes specified for the curriculum.
"Once you accept that people can rank and discriminate reliably without complete sets of distinct criteria, accepting grading - and using it when aggregating - becomes fairly straightforward." (Wolf, 1993).

Wolf regards the number of grades or marks that may be awarded on any given assessment, the way aggregation is done, or the preconditions or the amount of compensation that are allowed for an overall grade as "essentially practical decisions".

5.8 The question of obtaining both criterion-referenced information (to recognize the achievement of essential outcomes) and norm-referenced information (for purposes such as ranking).

Robert Glaser is regarded as one of the principal initiators of criterion-referenced assessment, outlined in a seminal paper in 1963. Glaser came to believe that criterion-referencing should

"...encourage the development of procedures whereby assessments of proficiency could be referred to stages along progressions of increasing competence". (Glaser, 1981, in Griffin, 1993)

Following Glaser, Griffin (1993) argues that the basic requirement for criterion-referenced assessment is the development of "progressions of increasing levels of competence", and he warns that "there is an enormous amount of work to be done in developing the continua and the need is urgent. Without the continua, there will almost certainly be a regression to the mastery learning movement of the 70's and an inevitable trivialising of the curriculum".

Drawing on Rasch's (1960, 1980) use of latent trait theory, Griffin suggests that "it is not necessary to assess and record performance on every task in order to understand each individual's progress. A person can be assessed by selecting tasks from the continuum to indicate the level of competence or proficiency reached by the student depending on which tasks are adequately performed." An important assumption underpinning this approach to competency-based assessment is that "a set of underlying continua can be constructed which describe development or growth in learning in specific domains of learning. The continua define constructs which are measurable, have direction and units of magnitude." (Griffin, 1993)

Griffin is not arguing a case for graded assessment. This model of criterion-referenced assessment may be viewed as being concerned primarily with whether or not a learner has achieved a particular level of competency or proficiency on a continuum; and this is determined by assessing whether or not the learner has adequately performed the tasks which are indicative of the achievement of that level.

Latent traits, as Griffin (1993) notes, are statistical constructs. As such they have several interesting properties. Tasks may be ordered in such a fashion that a continuum is constructed that will assist in understanding a trait (competency), for example, for a particular progression of tasks, it might be inferred that
"if a person can demonstrate competence on an uncommon or difficult task, it is likely that they will perform competently on the more common or less difficult tasks, and that this becomes more likely as the competency required becomes more common";

"if a person cannot demonstrate competence on a common task, then it is unlikely that competence will be demonstrated on the more uncommon tasks and that this becomes less likely as the competency required becomes less common."

Moreover:

"It is not necessary to assess and record performance on every task in order to understand each individual's progress. A person can be assessed by selecting tasks from the continuum to indicate the level of competence or proficiency reached by the student depending on which tasks are adequately performed. This is true criterion-referenced interpretation. It is necessary to identify the competencies, design the tasks that require those competencies and calibrate the continuum or progression as described by Glaser. 'Scoring a test' seems less relevant given that the performance is described in terms of nature of the ordered set of tasks performed rather than the count of the number of items correct or the number of tasks satisfactorily performed. The tasks then become the units of the measuring instrument, just like the units of length used to measure a physical object. This appears to be the implicit basis of competency-based assessment and should drive competency-based training and the measure."

Once a discussion of competency based assessment conceptualises competence in terms of a continuum, a complex progressive process of developing knowledge and skill, it is possible to develop a rationale for the recognition, in conjunction with criterion-referenced assessment, of a range of different levels of achievement or degrees of merit in competency-based courses, or in particular competency-based curriculum modules.

McGaw (1993) has proposed such a measurement model, based on the construction of an underlying scale. The scale would represent a continuum of tasks, related to a particular content domain, progressing from less difficult to more difficult along the scale. Performance criteria are defined for various points along the scale against which an individual's performance can be referenced. Individuals are assessed on test items or assessment tasks that correspond to particular points, particular levels of difficulty, on the scale. On this basis, an individual is assigned to a particular level of achievement on the scale. This is essentially a probabilistic model of assessment:

"the assessment process involves inferences from observations of actual performances on a set of items or tasks about how persons will respond to the population of such tasks, of which the ones represented are just a sample. There is never a single task that can capture all we wish to
know about a person's level of competence and there is no way of asserting with absolute certainty what a person's level of competence is;" 

"an individual's performance can be interpreted in terms of the underlying scale. The items below the student's location represent the kinds of things that the student can typically deal with successfully while those above represent the kinds of things the student cannot yet do."

This criterion referenced information may be supplemented with norm-referenced information, if this is needed for purposes such as selection:

"If the scale is applied to a representative sample of persons, the measurements of their performances will permit norm-referenced interpretations of individual performances."

That is, individuals in a group will be ranked in order, according to the locations to which each of them is assigned on the scale. The location on the scale of each individual will be based on his or her performance on test items or assessment tasks sequenced in increasing levels of difficulty.

McGaw (1993) considers that an approach to assessment that is wholly normative is "antithetical to the concerns of education" because it is "essentially competitive" and "focuses for individuals on how their performance stands in relation to those of others rather than on the nature of their performance". He recognises that ranking may be important for selection, but that it is "a distraction in the main game of education". Nevertheless, he recognises that, while criterion-referenced interpretations of individual performances may be more central to the educational endeavour, norm-referenced interpretations do have value and his assessment model permits them to be made if they are needed for purposes such as ranking.

McGaw illustrates his model primarily in the area of mathematics achievement in schools. For use in the vocational education and training sector, consideration would have to be given to the applicability of the model to a variety of vocational fields, to the number and variety of scales that would need to be constructed, and to the resource implications of the adoption of such a model.

The application of latent trait theory or item response theory to competency-based assessment for the purposes of "grading" is, however, immensely sophisticated mathematically, and could raise the question of the availability of the resources and/or the expertise that would be needed for an organisation like TAFE to implement it. (Dr Michael Bailey, personal communication)

McGaw does not believe that the different purposes for which assessments are used necessarily require different methods of assessment. He argues that the approach to
scale construction and assessment in his model enables valuable normative uses to be made of assessment data.

Similarly, Hager et al. (1994) discuss the use of "maps of increasingly competent performance" in conjunction with the assessor's "professional judgement". On such a "map" of competency,

"an individual is judged to sit somewhere on a continuum from 'not competent' to 'highly competent' on the basis of carefully established standards. These maps are derived from a holistic model which uses professional judgement but by reference to a continuum of performances."

The difficulties these authors identify with this approach include:

"the expense of developing the maps - presumably there would be a large number of domains for each occupation";

"the possibility that assessors would be constrained by the map and fail to see other evidence of high level performance."

The above discussion has focused on the application of latent trait theory to the design of an assessment system that would yield both criterion-referenced and norm-referenced assessment information.

A different question is whether or not it is feasible to derive norm-referenced information from criterion-referenced assessments. This question challenges the view that there is a dichotomy between these two assessment orientations.

Peddie (1993b) comments that there is a relationship between norm-referenced assessment and criterion-referenced assessment in that "standards are always set in large part by reference to what equates to norms of (expected) learner behaviour". The setting of levels of achievement would "correspond to similar expectations about what typical learners might find difficult". In this sense, according to Peddie, the emphasis in competency-based education on the nature of a skill and the specification of related assessment criteria is "less important than the conventional understanding of the norms of learner behaviour".

Similarly, McGaw (1993) argues that "the normative link was never really broken despite the claims of many practitioners" of criterion-referenced assessment:

"Course criteria cannot realistically be defined without some consideration of what is typically achieved by students in such courses."

Nevertheless, Peddie (1993b) suggests that in a well taught competency-based program it should not only be theoretically possible, but should also be a practical goal, for the outcomes specified for the program to be attainable by all learners. This extends to the setting of a merit standard:
"While it may well be the case, because of the underlying use of norms, that only a small group of learners does reach the merit standard, there should be a genuine attempt on the part of the standards-setting bodies to avoid a merit standard which deliberately attempts to engineer such a result."

Peddie, however, asks:

"whether those who are interested in a merit standard being developed would feel happy about a standard which turned out to be attainable by all learners in a unit."

In contrast, it is conventionally considered that criterion-referenced tests provide the best information about learners' performance in relation to the domain of content, and that norm-referenced tests provide the best information about individual learners' performance in relation to the group of learners. According to this view, it is possible to use assessment results of a criterion-referenced test to make norm-referenced interpretations, but this norm-referenced information is not optimal; similarly, it is possible to make "weak" criterion-referenced interpretations based on norm-referenced tests. It is argued, however, that assessment results on criterion-referenced tests may be relatively homogeneous, reducing their value for norm-referenced interpretation. Criterion-referenced assessment is not primarily designed to discriminate amongst learners. In contrast, norm-referenced tests are designed, for example by selecting test items of moderate difficulty and high discriminating power, to spread individuals out across a wide range of scores and to provide a valid and reliable ranking of learners. (Hambleton and Rogers, 1991; Nitko, 1984)

Hager et al. (1994) are concerned that if the achievement of essential module or course outcomes is determined by criterion-referenced assessment, and if relative merit is determined by norm-referenced assessment, this dual system "does not answer the question of what a merit learner can do that a satisfactory learner cannot."

An assessment system may, however, for different purposes, need to provide both norm-referenced and criterion-referenced information. Norm-referenced data may be needed for purposes such as selection. Nitko (1984) argues more generally that "norm-referenced data are needed to interpret fully an examinee's criterion-referenced test performance. ... Criterion-referencing and norm-referencing provide complementary information." The purpose of obtaining information about degrees of merit or levels of achievement may be not to determine what a merit learner can do that a satisfactory learner cannot, but to rank, in order of merit, learners who perform better than satisfactory learners. There would often not be any need to do this, but, where there is a need (for example, to facilitate selection), norm-referenced assessment is designed specifically to achieve that purpose directly. Such dual assessment of the same module or course content might not, however, be an efficient use of assessment resources.

It is argued that criterion-referencing and norm-referencing are best considered as polarities on a common spectrum, rather than as sharply dichotomised. In reality,
"purely" norm-referenced or "purely" criterion-referenced tests do not exist: rather, a given test might be considered to be more towards the criterion-referenced end of the spectrum (Dr Michael Bailey, personal communication) As discussed above, norm-referenced information can legitimately be derived from predominantly criterion-referenced tests; optimal norm-referenced information, however, would be derived from tests more towards the norm-referenced end of the spectrum.

In a competency-based system of education, criterion-referenced information about whether or not a learner has achieved the outcomes specified for a "pass" is of fundamental importance. It would not be acceptable to attempt to base such a judgement on the results of a norm-referenced test. If, however, additional, norm-referenced, information about learners is also required in particular circumstances, consideration needs to be given to the methods by which it should be obtained. One option for this appears to be to use norm-referenced assessment in addition to criterion-referenced assessment. Another option would be to use an assessment method which generates both criterion-referenced and norm-referenced information. This could involve the use of competency maps or competency continua, which may be based on the application of latent trait theory. Alternatively, it could involve the legitimate derivation of norm-referenced interpretations from assessments that are primarily criterion-referenced.

5.9 The form in which the results of "graded" assessment may be reported.

There is an interest in the literature in reporting through profiles. For example, Debling (1989) suggests the use of profiling of the "personal effectiveness attributes" of individuals, to provide "recruiting officers, whether in employment or education" with a basis for evaluating the potential of applicants for future success:

"Such personal effectiveness attributes are more normally profiled rather than pass/fail criteria being applied. Nevertheless there are criteria defined for each element of the profile, used in assessment and providing the basis for certification...."

"In particular in education, there is extensive interest in these academic, generic and personal transferable skills..... with respect to the personal transferable skills domain a profiling approach might be recommended."

Hager et al. (1994), however, while seeing the "use of profiles of achievement which detail demonstrated competencies (at various levels) in written form" as a method of "grading", identify a number of problems with this method:

- it is "time consuming for assessors";
- it is difficult for employers to interpret and "the evidence suggests that employers have no interest in trying to interpret such information";
- it is difficult for admissions officers in other educational systems to interpret, and in their case "aggregated grades are probably necessary".
Wolf (1995) agrees that "all the available evidence on employers' use of information" demonstrates that they want "a simple sifting device" rather than "detailed profile information", and this is especially so, "given the costs of recruitment, hiring, etc."

Much of the literature dealing with the recognition of one or more degrees of merit assumes the reporting of different levels of achievement in grades (such as "A", "B", "C", or "Distinction", "Credit", "Pass"), or perhaps in marks.

The research study carried out as part of this project aimed to identify stakeholders' preferences regarding the methods of reporting on levels of achievement.
6. METHOD FOR THE STUDY

In accordance with the project brief, a variety of techniques were employed to gather the required data.

Stakeholders were identified from whom it would be appropriate to obtain information:

- industry / employers;
- relevant government agencies;
- other education and training providers;
- Institutes / teachers;
- students;
- TAFE NSW Training Divisions.

An initial focus group consisted of members of a wide range of stakeholder groups, including TAFE and other providers, government agencies, and industry. The project brief tended to assume the availability of graded assessment as an option and focused on reporting methods in cases where graded assessment was implemented. In the initial focus group, however, some participants objected to any assumption that it was acceptable in principle to assess and report on the achievement of different degrees of merit within a competency-based education and training system.

Although the primary aim of this study was to determine stakeholder preferences for the form in which such assessments might be reported, it became important to provide participants in surveys and interviews with opportunities to state and argue their position on this underlying issue of principle, on which many respondents had strong views. This was consistent with the project brief, "to investigate what industry and others actually want from the reporting of assessment", and assisted the study to canvass a wide range of views on the reporting of assessment.

Similarly it was apparent that any discussion in the final project report of the forms in which the results of assessment might be reported should be accompanied by an explicit acknowledgment of this underlying issue and an exploration of the issue based on the relevant literature. (See Section 5 above, "The issue of 'graded assessment' in a system of competency-based education and training").

Following the initial focus group, information was gathered through interviews and questionnaires.

Multiple copies of questionnaires were sent to the State Managers of all 13 NSW TAFE Commission Training Divisions, responsible for curriculum development. Returns were received from all Training Divisions. Returns received from Training Divisions ranged from
a single consolidated response representing the Training Division's collective position, to a response from a staff meeting with the addition of a small number of individual responses, to individual responses only.

A questionnaire was sent to the Executive Officer (or equivalent) of twenty NSW Industry Training Advisory Bodies (ITABs). Responses were obtained from nineteen of the twenty, either by return of the completed questionnaire or through an interview.

In the case of private providers, data was gathered by interview.

In the case of TAFE NSW teachers, the method adopted was to survey the teaching sections in all colleges in a single Institute. The Hunter Institute of Technology was chosen because its colleges included both urban and country colleges. 180 questionnaires were distributed to heads of teaching sections. 107 responses were received.

Detailed information about participants in the study is provided in section 10, "Informants".

Data was not gathered from students because an adequate sampling of students across a range of fields of study and levels of course was not feasible within the resources available for the project. Rather than carrying out a token, and possibly misleading survey of students, it is suggested that a subsequent research activity could include the investigation of student reactions to these findings and to the issues raised in this report.
7. RESULTS OF THE STUDY

7.1 Assessment of levels of achievement in a competency-based system of education and training

Assessment may be concerned only with determining that a learner has achieved essential curriculum outcomes (such as achievement of specified learning outcomes, module purpose or course aim). In addition, assessment may be concerned with how well a learner performs in relation to such curriculum outcomes, that is, it may recognise different degrees of merit or levels of achievement.

The project sought to determine the extent to which such recognition of levels of achievement was considered acceptable in a competency-based approach to education and training.

Participants from Training Divisions were asked:

_Which of the following statements is closer to your own understanding of the principles underlying competency-based assessment?_

I) Assessing different levels of achievement in modules is not compatible with a competency-based approach to education and training.

II) Assessing different levels of achievement in modules is compatible with a competency-based approach to education and training, but is not necessarily desirable.

Statement (II) received greater support:

- all respondents from six of the thirteen Training Divisions supported statement (II) as did a majority of respondents in another five Training Divisions.

Other participants (industry, to TAFE teachers and to other providers) were asked:

_Which of the following statements is closer to your own understanding of the principles underlying competency-based assessment?_

I) levels of achievement should not be assessed in a competency-based system of education;

II) levels of achievement may be assessed in a competency-based system of education, but it is not necessarily desirable to do so;

III) levels of achievement should normally be assessed in a competency-based system of education.

Statements (I) and (II) each received substantial support amongst respondents in industry. There was relatively little support from industry respondents for statement (III).
TAFE teacher respondents supported either statement (II) or statement (III) in preference to statement (I), indicating a clear preference for recognition of different levels of achievement to be available at least as an option.

Most of the small number of other providers contacted supported statement (II).

A number of respondents endorsing statement (II) did so despite objecting to its inclusion of the words "but it is not necessarily desirable to do so".

A number of respondents made clear that they endorsed statement (I) because they believed that the principles underlying a competency-based approach to education precluded assessment of different levels of achievement, even though they themselves preferred assessment to recognise different levels of achievement.

The pattern of responses from different industry sectors tended to support the comments of individual informants that the reporting of different levels of achievement was likely to be seen as appropriate in sectors that had well-established systems of training and skills recognition, typically apprenticeship systems, such as automotive, building and construction, electrical, and printing.

Opposition to the reporting of different levels of achievement was likely to be associated with sectors, such as community services, textile clothing and footwear, and rural. In such industries there had historically been only limited provision for training and skills recognition, and there were large numbers of disadvantaged and/or unqualified people in the workforce (especially women, older workers and immigrants) who needed to be encouraged to seek training and recognition of their skills in association with award restructuring. In such cases, identifying degrees of merit was a separate issue to recognising the achievement of occupational competence, was perceived to have little if any value, and may in fact discourage workers from participation in training and assessment. For example, the following comment was provided by a national training liaison officer for a union in one of these industries:

"The Union believes that the system of reporting for competency based training should be 'pass/more training required' or in your words 'pass/fail' at ASF levels 1-3. It is important that obstacles or disincentives to training are not created by 'reporting on student's levels of achievement'. The Union is opposed to any training or assessment educational system which has the capacity to promote competition amongst employees and act as a disincentive for educationally disadvantaged workers to undertake further training. In competency based training an employee achieves the 'purpose' of the module which is based on the National Competency Standards. An employee is either competent or requires more training.... In [this] industry there has been little structured training for operators and production workers and the majority of employees are women from non-English speaking backgrounds. It is important that such employees are not threatened or..."
disadvantaged by assessment grading, so as we are all unified in the promotion of training and assessment."

The Training Division which provided courses in the industry referred to in this example considered that it would be inappropriate to have graded assessment in work-based operator level courses, but tended to keep the graded assessment option open in the case of technician level courses, the majority of which were institution-based.

A substantial minority of participants in the survey believed that it was not appropriate to assess different levels of achievement in competency-based education and training. Many of them declined to answer some or all of the subsequent survey questions that dealt with how different levels of achievement should be reported.

**In summary:**

The most widely supported view was that levels of achievement may be assessed in a competency-based system of education.

The implementation of this principle would require one or more methods of reporting on levels of achievement to be available as an option for use in appropriate circumstances.

### 7.2 Preferred method of reporting levels of achievement in final result notices for modules

Participants were asked:

*Which of the following ways of reporting levels of achievement in final result notices for modules would you wish to have available?*

- Grades
- Marks
- Verbal descriptions
- Other? (Please specify)
- Combinations of the above? (Please specify)

The most widely supported form of reporting, amongst both industry respondents and Training Division respondents, was grades, followed by verbal descriptions and then by marks.

Similarly, the most widely supported form of reporting, amongst TAFE teacher respondents was grades, followed by marks. The use of verbal descriptions received little support amongst TAFE teachers.
A minority of respondents wanted combinations of forms of reporting, but nearly all of the suggested combinations included grades (for example, grades and marks or grades and verbal descriptions).

One view was that in a competency-based approach to education, there should be an emphasis on reporting performance outcomes. It would be possible to specify learning outcomes in performance terms associated with each level of achievement to be recognised. These standardised statements could be issued with the student's result notice either instead of, or as well as, grades.

One view, applicable to this and many of the following issues, was that there were advantages in maintaining consistency over time, that there have already been many changes in the curriculum arena, and that there should be no unnecessary changes in the form in which assessment outcomes are reported in result notices.

In summary:

The most widely supported form for the reporting of levels of achievement on modules was grades.

There was widespread support for supplementing the reporting of grades with reports of verbal descriptions of achievement and/or marks.

7.3 The use of grades for reporting on the assessment of levels of achievement

(a) the preferred number of grades

Participants were asked to indicate, in the event that grades were used to report levels of achievement in modules, if they believed the number of grades should be fixed or variable across modules.

All participants were asked how many grades they believed should be reported if the number of grades was fixed.

Industry respondents and TAFE teachers, with few exceptions, supported a fixed number of grades.

Training Division respondents were more divided, but the more substantial and more widespread support amongst them was also for a fixed number of grades.

Many Training Division staff believed that the number of levels of achievement that it was possible to distinguish in a module depended on the nature of the subject matter or content of the module. This would tend to support making the number of grades reported variable across modules within courses. The main argument against this was that it would lead to confusion amongst students, employers and other clients.
On the question of the number of grades to report, industry respondents indicated support variously for two, three or four grades, but no single option emerged as having significantly more support than any other amongst industry respondents as a whole.

TAFE teacher respondents were divided in their support between three grades and four grades. There was very little support for two grades.

Amongst Training Divisions there was a clear preference for three grades, followed by two grades and then by four grades. Training Division staff were asked how many grades should be available for reporting if the number of grades reported was variable. Their clear preference was for a choice between two and three grades, although a choice between two, three and four grades received some support.

The small number of private provider respondents all preferred a fixed number of grades but there was no prevailing view on what that number should be.

In summary:

The survey found more widespread support for reporting levels of achievement through a fixed number of grades than for permitting the number of grades reported to be variable from module to module or from course to course.

The survey did not find a clear preference for a specific number of grades to be used at module level, except in the case of Training Divisions, where there was a preference for three grades.

(b) the preferred terminology for grades

All participants were asked to indicate a preference for the terminology in which to report grades. The terminology was illustrated in all surveys using three grades, and options given as follows:

A, B, C

Pass, Pass with Merit, Pass with Distinction

Pass, Credit, Distinction

other? (please specify)

In the case of Training Divisions, participants were asked to specify their preferences for grading terminology if two grades were reported, if three grades were reported and if four grades were reported.

There was no clear preference amongst industry respondents for any one kind of grading terminology.
There was no clear preference amongst TAFE teacher respondents overall for any one kind of grading terminology, but when analysed according to teaching areas different preferences did emerge. Support was mainly divided between A,B,C and Pass, Credit, Distinction.

The only clear preference amongst Training Division respondents was for the use of Pass, Pass with Merit in preference to A,B or A,C or other, if only two grades were used.

If three or four grades were used, the Training Division respondents as a whole did not indicate a clear preference for any of the terminology options available.

An advantage of A,B,C was that it was already in use and would not give rise to confusion over time.

An advantage suggested for Pass, Credit, Distinction was that it would provide consistency in reporting terminology between advanced TAFE courses and certain higher education courses.

An advantage of Pass, Pass with Merit, Pass with Distinction, was that it emphasised (in repeated use of the term "Pass") the most important aspect of assessment which was recognition that the individual had achieved the essential learning outcomes specified for the module.

Training Division staff were specially asked if they believed confusion was likely to occur if the term "Pass" was used in reporting assessment results for ungraded modules, as well as to refer to the lowest grade in graded modules. A majority of respondents in each of eight Training Divisions considered that confusion was likely to occur, but taken as a whole Training Division respondents were fairly evenly divided in their views on this question. It was suggested that result notices should make clear whether or not the module was graded, but making sure that clients note and understand the implications of this can be difficult.

Alternative terminology suggested by respondents included:

- in the case of two grades:
  - "Pass, Credit",
  - "Achieved, Achieved with Merit",
  - "Competent, Distinction",

- in the case of three grades:
  - "Competent, Merit, Distinction",
  - "Satisfactory, Good, Excellent"
In summary:

There was no clear preference for a specific kind of grading terminology, except where only two grades were reported.

In the case of two grades being reported, Training Division respondents had a clear preference for "Pass, Pass with Merit". Many Training Division staff, however, were concerned that confusion was likely to occur if the term "Pass" was used in reporting assessment results for ungraded modules, as well as to refer to the lowest grade in graded modules.

7.4 The use of marks for reporting on the assessment of levels of achievement

(a) the basis on which marks might be used

Participants from Training Divisions were asked:

Which of the following statements is closest to your understanding of the role of marks in competency-based assessment?

(I) marks are inappropriate in competency-based education and training, because it is not possible to make clear what a particular mark means in terms of knowledge and skills held by the student;

(II) marks may be used in competency-based education and training, if particular bands of marks can be shown to correspond to the possession of particular knowledge and skills;

(III) marks may be used in competency-based education and training, if they are accompanied by other information, such as grades based on stated criteria or verbal statements that clarify what knowledge and skills the student possesses.

Other participants (industry, TAFE teachers and other providers) were asked the same question, with the addition of a fourth statement:

(IV) marks may be used in competency-based education and training, provided that it is clear what mark represents achievement of the minimum or essential requirements of the competency-based program.

Training Division respondents were divided over whether or not the use of marks was appropriate in competency-based education and training: all respondents from five Training Divisions considered it appropriate; all respondents from four Training Divisions (and the great majority from a fifth) considered it inappropriate; respondents in each of the remaining three Training Divisions were divided.

Industry respondents were also divided over whether or not the use of marks was appropriate in competency-based education and training: for example, staff of 9 ITABs responded to this issue: responses from three ITABs considered marks inappropriate,
responses from four or five ITABs considered it appropriate, and responses from one ITAB were divided.

Amongst those Training Division respondents who considered the use of marks appropriate, no clear preference emerged for either statement (II) or statement III.

Amongst those industry respondents who considered the use of marks appropriate, no clear preference emerged for any of statements (II), (III) and (IV).

The small number of responses from private providers were consistent with the responses from Training Divisions and industry.

There was a clear preference amongst TAFE teacher respondents for the view that marks may be used in competency-based education and training, but no clear preference emerged for any of statements (II), (III) and (IV).

Against the use of marks, it was argued that if they were used, assessment schemes would remain largely unchanged and would be invalid in terms of criterion-referencing, because it was argued, they assume either norm-referencing or achievement on a performance continuum. How does a distinction between 67% and 68% relate to a criterion-referenced or a competency-based approach to assessment? This issue is quite significant because it was evident from survey returns that in those cases where assessment is reported in grades, or even in "Pass/Fail" terms, assessment results may initially be-recorded in marks and then converted for reporting purposes.

In summary:

The survey did not find decisive support for or against the view that the use of marks was appropriate in competency-based education and training, except in the case of TAFE teachers: there was a clear preference amongst TAFE teacher respondents for the view that marks may be used in competency-based education and training.

(b) the preferred range

Participants were asked:

*If marks are used to report levels of achievement in modules, what range of marks should be used?*

0-20?

0-100?

other? (please specify)

The overwhelming preference amongst Training Division, industry, TAFE teacher and private provider respondents was for a marks range of 0-100.
It was suggested by one respondent that holistic assessment based on module purpose lent itself to marking on a scale of 0-100 more than assessment of individual learning outcomes did.

In summary:

There was a clear preference for the use of a range of 0-100 marks, if marks were used.

(c) reporting on essential learning outcomes

Industry, TAFE teacher and private provider participants were asked:

*If percentage marks are used to report assessment results in a module, what minimum mark should represent achievement of a "Pass" or its equivalent?*

- 50%
- 70%
- 80%
- 100%
- other (please specify)

The mark should not necessarily be the same for all modules.

None of the options specified was clearly supported or clearly rejected by industry respondents as a whole. A minority of industry respondents, but three of the four private provider respondents, preferred that the mark should not necessarily be the same for all modules.

Of the 106 TAFE teacher respondents, 37 supported 50%, 19 supported 70% and 39 supported the view that the mark should not necessarily be the same for all modules.

In summary:

The surveys of industry, of TAFE teachers and of private providers did not find decisive support for fixing the mark representing "Pass" or its equivalent at 50%, 70%, 80% or 100%.

These surveys did not find decisive support for permitting the "Pass" mark to vary from module to module.

(This particular issue was not raised in the earlier survey of Training Divisions.)
7.5 The use of verbal descriptions for reporting on the assessment of levels of achievement

Participants were asked:

*If verbal descriptions were used to report levels of achievement in modules, what form should these take?*

(I) substantial detailed descriptions of the quality of the student's achievement by the teacher(s)?

(II) brief teacher comments?

(III) comments selected by the teacher(s) from a list of standard statements?

Respondents from ITABs preferred (II) or (III) over (I), and other industry respondents preferred (I) or (III) over (II).

No clear preference for any one approach emerged amongst industry respondents as a whole, however, or amongst private provider respondents.

Training Division and TAFE teacher respondents had a clear preference for (III).

One Training Division respondent suggested a system for generating detailed profiles for different levels of achievement: criteria for each level of achievement would be specified with reference to the module purpose; a grade would be recorded (but not reported) for each student based on these criteria; based on this grade a computer would generate an assessment report consisting of a detailed profile derived from the grading criteria.

In summary:

If verbal descriptions were used to report levels of achievement in modules, only Training Divisions and TAFE teachers demonstrated a clear preference for any one kind of verbal description: Training Divisions and TAFE teachers preferred comments selected by the teacher(s) from a list of standard statements.

7.6 Reporting levels of achievement in a course

(a) reporting on levels of achievement in a course: whether or not the option should be available

Participants were asked:

*Which of the following statements do you prefer?*

(I) overall levels of achievement should not be reported in awards for courses;

(II) it should be possible for overall levels of achievement to be reported in awards for some courses.
A large majority of TAFE teacher respondents preferred (II).

All respondents from six Training Divisions preferred (II). The respondents from each of the remaining seven Training Divisions were divided between (I) and (II).

Industry respondents were divided, but (II) was more strongly and more widely supported. This was also the case amongst private provider respondents.

One industry respondent argued that, while grading on modules against criteria for achievement in the module as a whole was acceptable, it was not acceptable to aggregate such module results to achieve a course grade. If this were done the grade, being an aggregation of unlike components, would not correspond to specified criteria. A course grade would be acceptable if there were criteria for achievement in the course as a whole, but this was not likely in the current policy context which encouraged increased flexibility within a course, implying the provision of numerous elective modules and course composition options.

In summary:

There was a clear preference that it should be possible for overall levels of achievement to be reported in awards for some courses.

(b) the preferred number of levels of achievement in a course to be reported

Participants were asked:

If levels of achievement in whole courses are reported, how many levels of achievement (excluding "Fail" or its equivalent) should be recognised?

two

three

other? (please specify)

There was a clear preference amongst industry respondents, TAFE teacher and private provider respondents for three levels of achievement to be recognised.

Amongst Training Division respondents, support was more evenly divided between two and three levels of achievement:

- all respondents from three Training Divisions preferred two,
- all respondents from four Training Divisions preferred three, and respondents
- from each of the remaining six Training Divisions were divided between two and three (or, in one case, two, three and four).
In summary:

There was a clear preference overall for the recognition of three levels of achievement at whole course level, although Training Divisions were more evenly divided between support for two levels and three levels.

(c) the preferred method for reporting on levels of achievement in a course

Participants were asked:

*If overall levels of achievement in a course are assessed, in what form should they be reported?*

- award grades
- other? (please specify)

[alternative forms to grades suggested earlier in the questionnaire were marks and verbal descriptions]

Few alternatives to award grades were proposed. The use of award grades was supported by a large majority of Training Division respondents, a large majority of TAFE teacher respondents, all private provider respondents and a large majority of industry respondents, including all respondents in ITABs.

In summary:

There was a clear preference for the use of award grades as the method of reporting levels of achievement in whole courses.

(d) the preferred terminology for reporting on levels of achievement in a course

Participants were asked:

*If awards are graded and two grades are used, how should they be reported?*

- Pass/ Merit
- other? (please specify):

*If awards are graded and three grades are used, how should they be reported?*

- Pass/ Credit/ Distinction
- other? (please specify)

[alternative terminology suggested earlier in the questionnaire were "A,B,C", "Pass, Pass with Merit, Pass with Distinction"]

Most respondents in all categories endorsed the use of the terms "Pass/Merit" for two grades, and "Pass/Credit/Distinction" for three grades.
In summary:

If two award grades were reported, the preferred nomenclature was "Pass, Pass with Merit".

If three award grades were reported, the preferred nomenclature was "Pass, Credit, Distinction".

(e) reporting of a grade point average for a course

Training Division staff and focus group participants were asked:

Would you support the calculation and reporting of grade point averages -

instead of award grades?

as well as award grades?

The reporting of grade point averages instead of award grades was rejected by the great majority of respondents.

The reporting of grade point averages as well as award grades received more support but was still rejected by most respondents.

In summary:

There was a clear preference not to report grade point averages, whether they were reported instead of award grades or as well as award grades.

7.7 Perceptions of teachers regarding student responses to the recognition of merit

(a) prospective students

The survey of teachers included the following question:

Do you believe that reporting different levels of achievement in a course/program, instead of pass/fail, could make the course/program more attractive to prospective students?

(b) the teaching/learning process

The survey of teachers included the following question:

Do you believe that reporting different levels of achievement in a course/program, instead of pass/fail, could benefit the teaching/learning process?
In summary:

The great majority of TAFE teachers in the Institute surveyed believe that reporting different levels of achievement in a course/program, instead of pass/fail, could make the course/program more attractive to prospective students. This was either the dominant or the unanimous view of respondents in all fields of study.

The great majority of TAFE teachers in the Institute surveyed believe that reporting different levels of achievement in a course/program, instead of pass/fail, could benefit the teaching/learning process. This was either the dominant or the unanimous view of respondents in all fields of study.

7.8 A note on the use of the terminology, "Pass/Fail"

Participants were asked:

In the case of a particular module, the aim of assessment may be only to determine whether or not a student has achieved learning outcomes that meet the purpose of the module. In this case, do you believe that the best form in which to report results is "Pass/Fail"?

The "YES" and "NO" responses to this question were both well supported. Many of those who responded "NO", however, wrote comments indicating that they were not focussing on the precise issue addressed in this question, that is, on the best form of reporting when the aim of assessment is "only to determine whether or not a student has achieved learning outcomes that meet the purpose of the module". They were clearly addressing a different issue: whether or not levels of merit should be recognised. Their rejection of "Pass/Fail" reporting was accompanied by arguments for grades, marks and the like. It could be that a number of "YES" responses were also focussed on this different issue.

Those respondents who addressed the question and opposed the use of "Pass, Fail" were especially concerned about the demoralising connotations of the term "Fail".

Alternative terminology suggested by respondents included:

- "learning outcomes achieved, learning outcomes not (yet) achieved";
- "competent, not (yet) competent";
- "acceptable, not acceptable at this stage".

A couple of responses from Training Division staff supported some form of profile reporting. Statements of specific module learning outcomes as stated in the curriculum, or improved statements of module purposes, could be the basis of such reporting, to make explicit in the assessment report what knowledge and skills the learner has demonstrated.
8. IMPLICATIONS

When a decision is taken to "grade" assessment in a module or a course, the benefits to be derived or purposes to be served need to be identified.

The number of levels of achievement or degrees of merit to be distinguished, and the form in which their achievement is to be reported, would desirably be consistent with the realisation of these purposes or benefits.

The recognition of merit could be catered for by the provision of one or two merit grades. Selection purposes might also be served by the provision of one or two merit grades, although more grades could serve this purpose more effectively.

The number of degrees of merit to be reported, and the form in which they would be reported (grades, marks, verbal descriptions) will, together with resource implications and feasibility considerations, influence the assessment method.

It is important not to underestimate the implications of what Wolf (1993) calls "essentially practical decisions" (see p.33 above). Deciding the number of grades to report is an example. Klauer (1984), in a discussion of different criterion-referenced grading models, comments that there are increased risks of inaccuracy when assigning multiple criterion-referenced grades as against making only a dichotomous judgment, and that it is important to ensure that grades that have been awarded either too high or too low "eventually balance themselves out". There is a paradox associated with increasing the number of grades that are reported: "the more grades that are used, the more the number of candidates misclassified rises but the more the severity of each misclassification falls: the fewer grades that are used, the fewer the number of misclassifications but the greater the severity of each misclassification so made. It is a neat situation..." (Willmott and Nuttall, 1975, in Cresswell, 1986) In his discussion of accuracy, or more precisely, reliability in awarding grades, Cresswell (1986) assumes that assessment aims to identify a learner's "true grade", and that what is reported (the "observed grade") will correspond to the "true grade" if assessment is reliable. There are then two different measures of reliability: the extent to which "observed grades" convey information about "true grades"; and the proportion of learners whose "observed grade" is the same as their "true grade". As the number of grades that can be awarded increases, these measures diverge: "observed grades" become more like "true grades", but a declining proportion of learners get their "true grades". The fewer the grades that can be awarded, the more severe are the grading errors. For this reason, if "grading" is to be used in internal assessments, for example to encourage learners to aspire to high levels of achievement or to provide feedback on the teaching/learning process, a large number of grades, or even percentage marks, may be preferable. On the other hand, for external reporting, a desire to ensure that there is a higher degree of correspondence between "true grade" and "observed grade", or a desire to report on the achievement of some form of grade-specific content-related criteria, might make it preferable to report in terms of only a few broad grading categories. If, however, external reporting of grades is intended to facilitate selection, increasing the number of grades will not reduce the efficiency of selection, and may for some procedures improve it. Some combination of approaches may also be possible, whereby broad grade categories are fixed based on performance-related criteria, and then
subdivided: the meaning attached to the broad grade categories would be primarily expressed in the achievement criteria, while the meaning attached to the subdivisions would be primarily an expression of relative merit. (Cresswell, 1986) Policy-makers and practitioners need sound technical advice when making such "practical decisions".

The question arises as to whether or not norm-referenced assessment information should be used to supplement criterion-referenced assessment information, if the aim of "grading" is to rank applicants for purposes such as selection.

Three broadly different approaches to "graded assessment" have been identified in the literature:

- the use of criterion-referenced assessment both to determine whether or not an individual has achieved the outcomes specified as essential for completion of a module or course, and to determine the degree of merit, or grade, the individual has achieved;
- the use of an assessment method that will provide both criterion-referenced information (to recognise the achievement of essential outcomes) and norm referenced information (for purposes such as ranking). This could involve:
  - the use of competency maps or competency continua, which may be based on the application of latent trait theory; or
  - the legitimate derivation of norm-referenced interpretations from assessments that are primarily criterion-referenced;
- the use of criterion-referenced assessment (to recognise the achievement of essential outcomes) and separate norm-referenced assessment (for purposes such as ranking).

There is a need for further research to explore the applicability and feasibility of these three approaches to "graded assessment".

Responses to the project's surveys revealed wide variation in the understanding of norm-referenced and criterion-referenced assessment and their relationship to competency-based education and training. In particular, it was evident that the use of marks is widespread within the vocational education and training sector despite the advent of competency-based assessment. "Pass/fail" judgments may be reported on the basis of recorded but unreported marks. Similarly, grades may be reported, based on bands of unreported marks. The extent to which the generation of these underlying marks is based on criterion-referenced, rather than norm-referenced, assessment is not clear.

It was suggested by some informants that while the primary purpose of final summative assessment is to determine "pass/fail" (or its equivalent), assessment during a program is often used to build confidence, to encourage and to motivate. If a program operates on the basis of cumulative assessment there may be a tendency for assessments to be distorted by the tension between the desire to encourage and the desire to report against a standard. This may have
implications both for the reporting of "pass/fail" judgments and for the reporting of "graded assessment".

There is a need for further research to explore current assessment practice, to determine the extent to which practitioners understand the principles and procedures associated with both criterion-referenced and norm-referenced assessment, and to consider ways in which current assessment practice might be informed and improved by an exploration of the three assessment approaches identified above.

The NSW TAFE Commission (1994) has identified a number of models for the calculation of grades for modules, and also for courses. For example, options for grading modules include:

- inference based on professional judgment with reference to module purpose;
- standard procedures for combining graded results within module;
- totalling and/or averaging marks.

Options for grading courses include:

- results from all modules may be combined to determine overall result;
- results from selected modules may be combined to determine overall result;
- averaged marks of modules contributing to award;
- standard procedures for combining graded module results within course;
- major piece of work assessing integrated outcomes of course.

Each of these options raises significant issues. For example, a problem with averaging grades, as Wolf (1995) points out, is that conflation of assessment information necessarily introduces "a degree of ambiguity about the underlying performance". If it is intended that a reported grade convey information about underlying performance, methods of arriving at that grade need to be evaluated. There is a need for research into the technical assessment issues involved in each of these options, to advise practitioners on the most appropriate assessment methods to use in particular circumstances for "grading" modules and courses.
9. REFERENCES


Byrne, Judy (1993), *The Option of Graded Assessment in Competency-Based Education and Training*, NSW TAFE Commission, Sydney.


Fletcher, Shirley (1992), *Competency-Based Assessment Techniques*, Kogan Page, London.


Griffin, P. (1993), "Issues in competency assessment: Don't revisit the seventies", in Testing Times: Conference Papers, TAFE NSW and NCVER, Adelaide


National Centre for Vocational Qualifications (1988), Assessment in National Vocational Qualifications, NCVQ Information Note No. 4. London

National Training Board, NTB Network, No. 15, November, 1994:


Peddie, Roger (1992), Beyond the Norm? An Introduction to Standards-based Assessment, New Zealand Qualifications Authority, Wellington

Peddie, Roger (1993a), Levels, Grades and Merit, notes prepared for participants in the workshop on levels, grades and merit in competency-based assessment, at the NCVER and TAFE NSW "Testing Times" Conference on assessment for competency based training, Sydney (3-5 November)

Peddie, Roger (1993b), Standards, Levels, Grades and 'Merit': A Critical Analysis, paper presented to the National Assessment Research Forum conducted by the Competency Based Training (CBT) Working Party of the Vocational Education Employment and Training Advisory Committee (VEETAC), Sydney

Training Agency (1988), "Assessment of Competence" (Guidance Note No. 5), *Development of Assessable Standards for National Certification*, Employment Department, Sheffield


Wolf, Alison (1995), *Competence-Based Assessment*, Open University Press, Buckingham

10. INFORMANTS

INdUSTRY TRAINING ADVISORY BODIES:

It is emphasised that the views expressed by the following participants do not necessarily represent an official position of the Boards or Councils of their organisations:

Ian Hetherington, Executive Officer, Arts Training NSW

Phillip Brown, Chief Executive Officer, Automotive Training Board (NSW)

Douglas Greening, Chief Executive Officer, NSW Building and Construction Industry Training Council Ltd

Dr Dennis Mather, Executive Director, NSW Chemical and Oil Industry Training Advisory Committee Ltd

Jill Gientzotis, NSW Community Services and Health Industry Training Advisory Body Inc

Leslie Mackay, Executive Officer, ACT Community Services and Health Industry Training Advisory Body

Bob Bebbington, Chief Executive, NSW Electrical and Electronics Industry Training Committee Ltd

Marcus d'Assumpcao, National Small Business Management Competency Standards Body

David Stone, Training Development Executive, NSW Fishing Industry Training Council Ltd

Jill Gillespie, Executive Officer, NSW Food Industry Training Council Ltd

Richard Stokes, Executive Director, NSW Furnishing Industry Training Council

Rod Vinten, General Manager, NSW Local Government Industry Training Committee Inc

Tom Stewart, State Coordinator, NSW Metals and Engineering Training Board

Trevor Davis, Executive Director, NSW Printing Industry Training Council

Elaine Jackson, NSW Retail and Wholesale Industry Training Council Ltd

Adrian Denyer, NSW Road Transport Industry Training Board Ltd

Richard McEvoy, Executive Officer, Rural Training Council of NSW Ltd
OTHER INDUSTRY RESPONDENTS

Sheila Beevours, Manager, NSW Business Traineeships, Hospitality Training Network

Robert Crabtree, Trade Consultant, Barry Rogers and Co, Pty Ltd

Senior Sergeant Debra Cussans, Acting Comm., Police Recruit Branch, NSW Police Service

R. Deegan, Director, Australian Customs Service

Max Dippert, MF Dippert Pty Ltd and the Toolmakers Association

Tony Fletcher, Management Consultant, Australia Post

Arthur Gagnuss, Human Resources Development Officer, BHP Steel

David Harries, Education and Development Officer, BHP Steel

Kevin Hummel, Partner, Training and Performance Solutions

Stephen Irving, Senior Engineer, Maintenance Education, BHP Steel

Richard Jenkins, National Adviser, Education and Training, Metal Trades Industry Association

Greg McSkimming, Risk Management Consultant, Kimberley Management A/Asia Pty Ltd

Mike Nacard, Manager, Group Training, Caltex Oil

Robert Parovel, Human Resources Development Officer, BHP Steel

Lyn Reid, SOGC, Joint Australian Public Service Training Council Secretariat, Public Service Commission

Bob Spence, Senior Manager, Computer Based Training, Commonwealth Bank

Rhonda Stewart, National Training Liaison Officer, Textile Clothing and Footwear Union of Australia

Margrit Stocker, Training and Development Consultant, BHP Steel

A group of ten Human Resource Development and Training practitioners from Commonwealth Public Service agencies.
A group of respondents surveyed by the NSW Electrical and Electronics Industry Training Committee Ltd on behalf of the project:

Graeme Barnes, Training and Development Manager, Matsushita Electrical Co (Aust) Pty Ltd

Colin Blundell, Resource Manager, Boral Elevators

S.J. Byles, Technical Services Manager, PCCC Electricity

Alun Gillies, Administrative Officer, Central West Electricity

Stephen Goode, Stephen Goode Electrical

Rod Gunton and Gordon Strachan, Coordinators, Apprenticeships and Curriculum Development, Prospect Electricity

Barry Hails, Engineering Officer, Murrumbidgee Electricity

Peter Halverson, Training and Development Manager, Orion Energy

Hany Hanna, Manager, Electrical Engineering, Southern Riverina Electricity and Water

Cindy Hastings, Acting Training Officer, Northern Rivers Electricity

Fred Head and Darren Jensen, Group Training Manager, National Electrical Contractors Association

E.B. Isaacs, Secretary/Treasurer, Electrical Apparatus Service Association

Derek Lark, Director, Network, Tumut River Electricity

Peter Longhurst, Manager, Network Operations, Monaro Electricity

Ian MacDonald, Admin. Officer (Resources), Ulan Electricity

George McCready, Human Resources Manager, Western Power

Allan McInnes, AGM, and Wayne McMillan, Safety Training Officer, Namoi Valley Electricity

Michelle Miran, General Manager, Illawarra Skills Development Ltd

J. Moffat, Technical Officer, and R. Wemyss, Workshop Superintendent, North West Electricity

Kevin O'Shea, President, Refrigeration Air Conditioning Contractors Association
Greg Paterson, Safety and Training, New England Electricity

Graham Pickup, Safety/Training Coordinator, Southern Mitchell Electricity

Ken Pidcock, Depot Superintendent, and Geoff Hennessey, Safety and Training Officer, Ophir Electricity

Franz Venhaus, Personnel Manager, Mitsubishi Electric (Aust) Pty Ltd

Judith Wright, Staff Development and Training Manager, Illawarra Electricity

OTHER PROVIDERS AND RELATED AGENCIES

Jan Andrews, Director of Educational Services, College of Nursing

Eileen Davis, Director of Studies, Australian College of Applied Psychology

Lawrence Hanley, Principal Lecturer, Sydney Graphics College

Trevor Lloyd, Director, Training, HQ Australian Defence Forces

Ian Maddock, Staff Officer, Training Projects, Directorate of Individual Training and Educational Policy, HQ Australian Defence Forces

Ron Perry, Director, Institute of Counselling

Sandra Tweedie, General Manager, NSW Vocational Education and Training Accreditation Board

John Williams, Senior Research Officer, NSW Vocational Education and Training Accreditation Board

Laurier Williams, Managing Director, and Kamora Chhan, Academic Consultant, Williams Business College

TRAINING DIVISIONS, TAFE NSW

Arts and Media

Tom Beggs
James Black
Dick Cooper
Alex Dalrymple
Mehrassa Farjad
Dick Middleton
Peter Wilde
Basic Work Skills

Lynda Henderson (on behalf of staff meeting)
Mary O'Sullivan
Neil Vickers

Building and Construction

Argy Beletich
Evalina Chan
Keith Cullen
Mick Hall
Christine Pracy

Business Services

Jeannine Angilley
Brian Burnett
Bob Burr
Tom Clemens
Pam Collier
Tim Dein
John Irvine
Lyn McRae
Kevin Reid
Margaret Summerfield
Frank Tweedie
Bev Wilson
C. Zanardo

Engineering Services

Brian Thomas (on behalf of the Training Division)

Foundation Studies

Ian Fegent
Maire Sheehan
Diane Warwick

Information Technology

Collective response from staff meeting
Focus group response from Branko Kulevski
Manufacturing
Sue Boswell
Lee Cummings
David Lewis
Julia Nancarrow
Peter Sheedy
Brian Syms
Raju Varanasi

Personal and Community Services
Shirley Dungate
Kon Medway
Gillian Morris
David Wallace
Gary Wood

Pre-Vocational Programs
Ros McLean
Kate Mudie

Rural and Mining
Alan Fridley

Tourism and Hospitality
Frances Farmer (on behalf of ten Training Division members)

Transport
Tom Nolles
Barry Palmer
Ross Parton
Terry Saunby

TEACHERS, TAFE NSW
Teaching sections of the Colleges in the Hunter Institute of Technology
OTHER TAFE RESPONDENTS

Rosemary Livingstone, Project Manager, Key Competencies Project

Steve Parkinson, Project Officer, Workplace Learning Project

Megan Ramadge, Training Consultant, Hunter Institute

Linda Simon, Secretary, TAFETA, NSW Teachers Federation

OTHER INFORMANTS

Mike Bailey, Faculty of Education, University of Sydney

Geof Hawke

Nick Lewocki, Labor Council of NSW

Tom Ruben, Executive Officer, Faculty of Medicine, University of Sydney

Roger Wescombe, Legal Practitioners Admission Board

Hugh Wilson, Faculty Secretary, Faculty of Dentistry, University of Sydney
11. APPENDIX

The following short paper by Ros McLean, Principal Officer, Joint Secondary Schools TAFE, Pre-Vocational Training Division, February 1994, details current issues to the assessment and reporting of vocational courses (made up exclusively of TAFE subjects) included in the Higher School Certificate:

Introduction

The Joint Secondary Schools TAFE Program allows senior secondary students to undertake some of their HSC Program of study in TAFE. JSSTAFE courses are provided on a fee for service basis to government schools, funded by the Department of School Education, and to non-government schools. It is the decision of the non-government school system authority as to whether the cost of participation in the program will be met out of school or authority funds or passed on to individual students. In 1995 approximately 20,000 students will participate in the JSSTAFE program.

Types of courses

JSSTAFE courses are either Board-developed or Board-endorsed.

Board developed

There are three JSSTAFE Board-developed courses, Accounting, Travel and Electronics Technology (2 year 2 unit). These courses count towards the HSC, attract a TAFE Certificate and are eligible for inclusion in the calculation of the TER. As such, a mark derived from assessments in all component subjects across the two years must be reported to the Board of Studies. In 1995, 3000 students are undertaking Board-developed JSSTAFE courses. A small number of these students are TAME HSC Pathway students.

Board-endorsed

Board-endorsed courses count towards the HSC but are not eligible for inclusion in the calculation of the TER. One and 2 unit courses, designed to meet the requirements of the Board of Studies, are made up of subjects drawn from major award TAFE courses.

Reporting

Up until 1994 where competency based courses which report only pass/fail were included as part of a JSSTAFE course (eg Metals and Engineering modules), students were awarded 30 marks for a fail and 70 marks for a pass, for the purposes of reporting to the Board of Studies. This was not an ideal situation but the only one available under the circumstances.

From 1995 the Board has advised that a mark will not be required for Board-endorsed JSSTAFE courses thus eliminating the less than ideal situation described above. The Preliminary (Year 11) record of Achievement will report satisfactory or unsatisfactory performance. The HSC Record of Achievement will report satisfactory or unsatisfactory performance and will refer to separate TAFE Result notices.
These result notices will be used to show evidence of the student's performance in each subject in the JSST AFE course in order to gain entry to higher level TAFE courses, other courses and/or to gain employment. They must therefore be reported in a form that is acceptable and meaningful to TAFE, other providers and employers.

**Board-developed courses**

For Board-developed courses, however, a mark will continue to be required by the Board and the universities.

The previous practice described above for Board-endorsed courses which contain pass/fail subjects is totally inappropriate for courses which contribute to the TER. To apply this practice to Board-developed courses would be most unjust.

This will have to be taken into consideration in the reaccreditation of the Travel and Electronics Technology courses which are due to be reaccredited this year. (The Accounting Certificate was reaccredited in 1994 in CBT format but still allows for the generation of a course mark)

The issue of reporting will also have to be taken into consideration in the development of more Board-developed courses made up exclusively of TAFE subjects which is planned for 1995. Like the existing three Board-developed JSST AFE courses these courses will also be available to TAFE HSC Pathway students. The development of new Board-developed courses will be severely restricted if TAFE adopts a pass/fail approach to reporting. Even a graded approach (with a limited range of marks) would not accommodate the Board of Studies and universities reporting requirements.

The development of competency based courses/modules which report only pass/fail clearly presents a problem if courses are to be included as part of the HSC and included in the TER calculation.

While this is a problem associated with the bringing together of two different types of assessment philosophies (and one that is clearly much broader than just the method of reporting) it is nevertheless a significant issue which must be considered as part of TAFE's deliberations on reporting of achievement.

If it is not considered the needs of multiple stakeholders - DSE, the non-government secondary sector, the Board of Studies, TAFE, the universities and students who wish to include a Board-developed vocational course(s) in their program of HSC study, are unlikely to be met. If these needs are not met weight will be added to existing perceptions that vocational courses are of lower status only relevant to the less 'academic' students not wishing to pursue university studies. This is surely contrary to the underlying principles of the national training agenda.

Perhaps at issue is also the method of recording assessment outcomes as distinct from the method of reporting to students/employers. At this point in time and for the foreseeable future the ability to derive a mark for reporting to the Board of Studies is critical if vocational courses are to take their place in NSW as an integral component of HSC offerings which contribute to university entrance.