Introduction

The purpose of this paper is to discuss procedures that lead to a differentiating score for Vocational Education and Training (VET) subjects at the Year 12 level, giving them parity with the more traditional subjects used for university selection. It arises from two other projects. One examines the procedures and outcomes involved in the synthesis of multiple sources of data into a single score and the subsequent interpretation in terms of an inference of competence as reported at this conference by Woods and Cooper (2002) and by Connally and Jorgensen (2002). The Australian Research Council funded that study and we refer to it in this paper as the 'competency synthesis project'. The second project investigated ways in which Year 12 students could be assessed for tertiary selection regardless of the nature of subjects studied and it is referred to the 'Year 12 project'. In the Year 12 project, vocational subjects that form part of the senior secondary curriculum and are linked to the Australian Recognition Framework (ARF). They could be used in the calculation of tertiary entrance scores and be treated as equivalent to the other subjects generally used for this purpose. The Australian National Training Authority (ANTA) funded the Year 12 study. Both projects shared the same theoretical basis that links scaling techniques to a process of synthesis based on multiple data sources.

The Year 12 project

At present Year 12 VET subjects are treated in a range of ways for purposes of ranking students for university selection. They are variously incorporated into other subjects, not used at all, or given the status of a 'bonus points' addition to the tertiary entrance score. Part from some pilot projects they have, in general, not been used in any way comparable to the 'traditional' university entrance subjects. We investigated ways in which study units that were assessed on a dichotomy of 'competent'/non competent' could yield a differentiating score, which could in turn be used by tertiary entrance bodies in a scaling or other procedure to rank students for tertiary selection. It was not an attempt to incorporate grading into the
assessment. A full description of how Year 12 Vocational subjects are assessed and reported is contained in the project report (Griffin, Gillis, Keating and Fennessy, 2001). In pursuing the possible Year 12 Vocational subject differentiating score, there were a number of conditions that had to be met. The general approach had to be consistent with the requirements of industry training packages; the differentiating scores had to be treated in an equivalent manner to core subjects by tertiary entrance authorities; parity of esteem between VET and Core subjects had to be maintained; a safety net was needed for the student group for whom the VET subjects were originally developed; and the assessment workload of both students and staff needed to be protected and minimised. The assessment outcomes had to be able to satisfy both the reporting requirements of the Australian Recognition Framework as well as those of senior secondary certificates. Hence assessment tasks were required such that they could produce data that could be interpreted and reported for two different, yet complimentary purposes, these being for 'selection' and 'recognition'. Another important distinction also needs to be made. These issues are not so much about the assessment process but about the way in which the outcomes of the assessment is reported. Much of the effort therefore is put into preparing records of observation for a range of reporting formats. In a recognition system, an insistence on a dichotomy of 'competent' / 'not yet competent' tends to pressure the recording system to only record the 'can/cannot' data and to then force assessment task design into a mode that only allows this form of recording. This paper argues that opening the recording process to the option of differentiating and recording levels of quality in observed performance, provides new opportunities for the assessment task design, recording procedures and in particular for the possible inferences and the reporting process. Any process of altering assessments will always have implications for each of these stages of assessment and reporting. It is not possible to change only one part of any assessment procedure without affecting each of the other stages of observe, record, infer and report.

**The competency synthesis project**

In the competency synthesis project, ways that assessment data can be combined from a number of sources into a single decision framework were sought and this has been reported by Connally and Jorgensen (2002). The study evolved out of a need to design, develop and validate strategies to synthesise multiple sources of evidence to make judgements associated with higher order competencies that could be directly observed nor simulated. The methodology merges two approaches to performance based assessment. The first encapsulates features of a 360-degree feedback assessment model that is widely used in performance appraisal and performance management assessment that is largely observation based, and includes other forms of evidence such as trainer/assessor judgements of portfolio task performance and paper based tests. The second approach uses psychometric methods that allow a synthesis of data types into a single measure while controlling for extraneous influences such as observer stringency. The methodology and the instruments have been designed to meet the requirements of the industry competency standards for public safety and public service for credentialling purposes.

**Comparability of parity and intellectual demands**

Comparability of standards associated with assessment are associated with parity of esteem and the intellectual demand of assessments and subjects. With a consistent approach to the development of tertiary entrance scores, parity of esteem should not affect the students' chance of selection. If however, the VET subjects are not on parity with other subjects, in terms of intellectual demand, then a lack of comparability of intellectual standards may diminish a student's chance of university selection. Part of the role of VET year 12 subjects is an encouragement of less academic students to remain at school. However, in order to obtain parity of esteem and intellectual demand there is a perceived need to attract the more
academically able students into those subjects. If this procedure is overtly successful, the less able students will be dissuaded from enrolment into VET subjects, but the subject will have achieved parity with the general subjects. The parity paradox is that by pursuing esteem, one of the goals of the VET subjects will be lost in that the very group of students that was expected to enrol may opt not to. As a consequence of the pursuit of parity, comparability of standards needs to be carefully handled. If the solution to this is a type of streaming of VET subjects, with lower demanding subjects catering for lower capability students, then the goal of comparability of standards is also compromised.

Every ACACA and education and training agencies across Australia has been consulted as well as a selection of RTOs and industry training advisory bodies, and tertiary admissions centres. An analysis of the assessment procedures used by the ACACA agencies in subjects such as music, drama, art, where performance assessments and an inference of competence are used were also examined. Similarly, the procedures that were based on performance or product that could be incorporated into competency-based assessment were examined.

Three aspects of assessment that could affect the decision making process needed to be taken into account. These are...

1. Where the task was developed;
2. Where the task was administered.
3. Where the assessment was scored (judged);

In addition to this there were also two main assessment paradigms referred to above.

1. **Selection** and **promotion** procedures are needed when applications for places exceed the number of places available. Differentiation among applicants is needed under these circumstances and one way of achieving this is to recognise differences in the quality of outcomes, performances and products produced by the applicants.

2. **Recognition** implies that there is no place for differentiation or ranking. This purpose simply needs to give due recognition for having acquired competencies, completed a course or attended a program.

In addition to the assessment paradigms, three broad areas of concern were identified. These were associated with a lack of emphasis on quality of performance, the quality of assessment task development needed for this and a serious lack of quality control over interpretation of evidence. This led to the following recommendations.

1. The use of a Standards Referenced Framework.
2. Assessment task design that recognises quality of learning and performance.
3. Quality control over various influences on judgement in a competency context.

**Standards Referencing**

A standards referenced framework is a subset of a criterion referenced interpretation approach to assessment. Criterion referencing is the process of monitoring and interpreting progress along a continuum of increasing competence. Glaser (1963) restated his original definition in 1981. Standards’ referencing is a type of criterion referencing in that levels (or bands) are defined along the progression for interpretive purposes. Grading and criterion
referencing are not the same thing. Grading is not a characteristic of criterion referencing and is not usually associated with the standards referencing. Grades are indicators of relative performance, usually within a normative framework, and cannot be interpreted as explicit statements of standards unless the standards-referenced system exists a priori. Reporting requirements for a competency system require a description of capability against the training package standards. Reporting for ACACA agencies require a differentiating score that can be scaled for university selection purposes. These joint requirements can be satisfied within a standards referenced system and hence this is the basis of our argument for such an interpretive approach. If we start with the description of capability it is possible to change the qualitative data to a quantitative summary, but it is difficult to reverse that process. If we start with the grade, it is difficult to provide the explicit description of capability. This is especially true if the grades are based on a normative distribution. They are only labels for ordered categories of students and have little to offer for finer differentiation and less to competency recognition. A norm-referenced system is not essential to rank students. Students can be ranked in both a criterion-referenced system, as interpreted within the ARF, and a standards-referenced system as interpreted by ACACA agencies.

Typically, in vocational assessment, a performance criterion is defined as a task to be performed and, usually, the assessment does not involve specific reference to the quality of the performance. Moreover it is expected that all performance criteria be satisfied before a decision of competence can be made. The problems associated with this approach were widely published in the experience in the 1970s (Griffin, 1993; 1995; Wolf, 1993). It was argued that this approach to assessment reduced the assessment to a level of trivia that had minimal, if any benefits to the stakeholders (e.g., students, schools, universities, employers, teachers). Yet we continue to pursue this approach to assessment in a CBA model.

Competency assessments focus on two levels only (competent / not competent), (recognition) while educational accreditation frameworks require the use and reporting of scores (selection). If criterion referenced continua of increasing competence (Glaser, 1981) can be developed and agreed levels on these continua used for defining scores for the senior certificates, an agreed cut point of competent/not yet competent can guide decisions for the Australian Qualifications Framework. This would aid the inclusion of VET in School subjects in the development of tertiary entrance scores. If a standards referenced approach to assessment is adopted, a decision needs to be made about the most appropriate way in which to express the criteria.

By focussing on the quality of performance, emphasis is placed on criteria that reflect how well the task was performed as well as a consideration of the knowledge and skills that underpin the tasks. RTOs that have implemented a graded competency based system, have tended to distinguish between levels of performance according to generic criteria that relate to the task.

Another approach involves the use of criteria that are expressed in terms of relative differences in the knowledge, skills and understanding that need to be demonstrated. In most instances, assessors or teachers judge performance and record their judgements on a rating scale format. Several ACACA agencies use comparative terms (e.g., limited/sound/thorough understanding) to differentiate between the levels of performance in VET study units. These are translated onto a rating scale and scores are determined by either tallying marks or averaging the ratings.

The use of comparative terms can, however, create problems associated with consistency of interpretation and hence reliability. Where this approach is used, exemplar materials need to be made available to teachers / assessors, and additional training is required to achieve a
common understanding of the terms. Both the ACT and NSW Boards of Studies, for example, are planning to avoid the use of such comparative terms and move to more specific criteria.

A standards referenced approach makes use of criteria that specify varying levels of quality performance expressed in the form of ordered descriptions of learning outcomes and competency. That is, it is content specific and describes varying levels of skills, knowledge and understanding underpinning the competencies. This approach is consistent with both the NSW and WA proposed standards-referenced model for reporting student outcomes and is an extension of the Victorian, Tasmanian and South Australian approaches. It requires the development of 'field of study' 'training package' specific continua of increasing competencies. Holistic judgements of competence levels are then used through the synthesis and integration of evidence. Assessment tasks (with varying levels of difficulty/complexity) and students (with varying levels of performance) can be mapped on to a single scale. This approach allows an empirical development and validation of broad bands of competencies for each subject and provides subject scores for each student. Criteria, within a criterion referenced assessment framework are thresholds on a continuum and as such they enable us to interpret evidence (indicators) and locate a person or a performance on the continuum. Collections of those indicators become a band level within a standards-referenced framework and this in turn helps to interpret competencies in a referencing system. There are several steps in the complete process of defining the levels of competence and while it appears that the process is linear, it is in fact somewhat iterative and the levels presented in this paper are illustrative only. Verification is always needed.

1. The Qualitative dimension - a single dimension?

The first step should always be the preliminary definition of the variable sometimes referred to as the qualitative dimension underpinning the development of competence. That is establishing a hypothesis of what is being assessed. What does the high end look like? What does the low end look like? Can intermediate levels be proposed. Clearly this is a step that is done by specialists in the area. It is important to question whether there is only one dimension to the competence. That is, is may be that a unit of competence demands more than a single competent decision. It may be there more than one competence underpins the unit. It is important to realise that this has never been tested. There is an implicit assumption in the training package series that each unit is underpinned by a single dimensional skill or competence. This series of levels and the overall proposed continuum becomes the basis for a final validation of the empirically defined process after trials. Validity is supported if the final interpretation of the continuum after trials and data analysis supports the initial proposition. This cyclic approach enables the developer to propose a continuum and then constrict tasks to measure performance of persons on that continuum.

2. Designing (sub)tasks to fit the defined dimension

Specialists in the competency unit need to identify the appropriate assessment tasks and sub tasks and to spread these tasks in terms of difficulty over the range of the dimension.

3. Defining rubrics

Specialists need to be involved in the definition of the expected levels of quality in the performances. These levels of performance quality are then used as the scoring rubrics. If empirical analysis of this is not going to be used, it is essential that the specialists also go through the exercise of placing the rubrics on the grid.
4. Trial and analysis

Once the rubrics (indicators) are identified their relative positions on the continuum need to be determined. This is achieved in two possible ways. Where the team of specialists is confident that they can accurately place the indicators on a grid as shown above, without data, this process can be undertaken. This consensus moderation procedure of defining the relative levels of performance quality can lead to stable placement (Bateman, 2002).

A second approach is to make observations of candidates and complete a rating sheet such as the one shown in Figure 6. The scores or relative positions assigned to these indicators can be used to calibrate the indicators and place them on a continuum using item response, modelling approaches such as reported by Connally and Jorgensen (2002). Where this is done, it can be simplified if a checklist is used and every indicator is recorded as present or absent. This dichotomous approach enables each indicator to be mapped onto the continuum. If the indicators are recorded using a rating scale where the location represents a score of one then a second indicates a score of two or more rather than one then a score of three or more and so on. Placing these indicators on a continuum of difficulty (or of competence demanded) enables a direct interpretation of the continuum to be established using a partial credit (masters, 2002) analysis.

5. Setting Levels

This is possibly the most difficult step. It involves placing cut points on the continuum in such a way that the levels of competence are interpretable, separable, and distinct. Wilson (1999) calls them 'segments' and the process 'segmentation of the variable'. Griffin (1986; 1990) has always called them 'levels' or 'bands' and the process 'profiling'. It involves several interrelated steps. If an empirical analysis has been used, an examination of clusters along the continuum is often the preliminary step. If these clusters are well separated (or segmentation is clear) then an examination of the rubrics and their clustering behaviour can help to interpret the levels of competence. In addition, guidance might be obtained by calculating the number of item strata (Wright and Masters, 1982), although this may not bear any relationship to the number of clusters.

6. Interpreting the Continuum

When clusters are agreed on and overall levels are identified and interpreted it is possible to gain an overall interpretation of the competency itself. The variable (or underlying continuum) is a synthesis of the levels. The cohesive nature of these levels gives meaning to the overall continuum.

7. Refining the continuum.

Sometimes, individual indicators don't seem to fit with the overall interpretation of the continuum. Decisions need to be made as to whether these isolated indicators need to be revised in terms of their expression in the light of the interpretation of the other indicators. Alternatively new indicators might be needed to fill gaps. It is also possible that some indicators are so different to others in the series that they detract from the overall interpretation of the level or the continuum. These might be removed. There have been instances where the indicator seems to fit better in another level. This requires some examination of the sample size, measurement errors and some judgement as to whether the indicator should be included at all. These actions should of course be non-extensive. If the continuum requires a great deal of revision, there may be something wrong with it initial definition or with the tasks or rubrics and a decision may be needed regarding the competency and its assessment.
8. Validating.

The overall interpretation of the level and their cohesive nature against the initially proposed variable and its levels needs to be validated. Where this is substantially matched validity is enhance. Once this is achieved is should be possible to select one of the levels as the competency level. In training packages this should match closely to the collection of performance indicators used to define competence. This advice however makes many assumptions about the standards, the training package and the dimensionality of the assessment. Where there is a serious mismatch between the indicators and the continuum, specialists in the competency domain may need to deliberate on the overall validity of the competency unit, the assessment regime, the interpretation or all of these. This dilemma remains the largest issue for this kind of empirical analysis o competency. The training package units of competence are endorsed as valid, but the developmental aspects of the levels have never been tested. If the match is acceptable however, the schema depicted below can serve as a workable framework for competency assessment.

Interpretation of an assessment is weakened if it is based on broad, generic criteria. Comparative terms like "basic', 'good', 'established', 'superior', 'better' and so on require interpretation and make consensus or exemplar moderation essential. However if these are replaced with directly observable criteria, the reliability and consistency of the observations and interpretations will improve. The criteria need to be content specific describing the varying levels of skills, knowledge and understanding that underpin the competencies of interest. Within a standards-referenced framework, there is no need to report students' learning in terms of grades. There is, however, a need to distinguish between levels of performance that are defined within a standards-referenced system. The levels need expert knowledge to develop and interpret.

Obviously those responsible for this definition of the levels must have the appropriate background and expertise. Levels of performance are however independent of the distribution of scores or of other ways of recording the quality of performances. The levels of performance are best defined within a standards-referenced framework. However, a shift to this framework will not, in and of itself, solve the need for raw data that can be used in developing differentiating scores. Acknowledgment that some tasks and subtasks are more difficult than others is necessary if differentiating scores are to reflect differentiating performances.

Allowing a mix of reporting modes in a standards-referenced system might resolve this dilemma. Possibilities include the use of a competency dichotomy as practiced within a CBA framework and a standards-referenced framework which can accommodate a range of levels as well as each of the year 12 report formats that are currently used in states and territories at the senior secondary level. Such a framework can also meet the needs of a range of stakeholders. It would be possible, for example, to provide reports based on strand descriptors (areas within subjects or units), completed units of competency, or even strands within units. Such an approach would also be possible to report in terms of competent/not yet competent dichotomies, grades or a differentiating score.

Task Development

Task development is also an area in which expert input is required pertinent to training package and workplace assessment expertise. The main concern regarding assessment design is the emphasis on checklists of tasks performed and a lack of emphasis on the other aspects of competency such as job role, contingency, transfer and linkage to other tasks. When the lack of concern for quality of learning is added to this the current assessment task design in vocational assessment was regarded as unsuitable for differentiation purposes.
The Victorian Certificate of Education (VCE) system has recognised this and in retail, hospitality and electrotechnology industries, has added emphases on underpinning knowledge and introduced central examinations, but has completely separated the decisions and assessments for competence and for selection resulting in a requirement of competence as an a priori condition for central assessment.

School and/or workplace--based assessment is used in all states and territories for vocational education. A wider range of course objectives/learning outcomes could be assessed if school-based assessments were combined with central examinations. In an indirect way, reliability could also be increased as repeated 'measures' of students' performances could be obtained over a period of time (as opposed to a one-off assessment).

However, there is a trade-off associated with school based assessments in both reliability and validity (McGaw, 1997). The use of school-based and/or RTO-based assessments introduces other threats to validity and reliability (refer to Messick 1989; 1992; 1994; Cronbach 1971). They are not necessarily fairer than external examinations because they can introduce other forms of bias; particularly those associated with judgement error (Gillis, Griffin, Trembath & Ling 1997). The authenticity of the evidence produced by the students is open to question, as was the case in assessment systems where independent student projects were required in high stakes common assessment tasks. To control for this, most Australian states and territories have implemented some form of moderation process to ensure comparability across schools and to follow up in matters of authenticity.

Face validity (or perceived credibility) is important. The use of performance or product based assessments, whether administered by the school and/or RTO should be a central feature of competency based assessment. This is emphasised by training package advice that the assessments should occur in either real or simulated workplace settings. Given this, and the need for assessment data to differentiate across varying levels of performance within VET courses, a number of system level decisions need to be made regarding responsibility for designing, administering, scoring and reporting assessments. Without assessment data the grids that are illustrated in Figure 1 serve as an approach to linking the rubrics to the hypothesised underlying continuum.

Unit name: Design and manage activities which exercise elements of emergency management

Unit Code: PUAEMR006A

Unit Descriptors This unit covers the competency to design, manage and evaluate activities which exercise elements of Emergency management

DRAFT ONLY

<table>
<thead>
<tr>
<th>Item</th>
<th>Performance Quality Level</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Low</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>The need for exercise is identified in consultation with stakeholders</th>
<th>Identify which risk management strategies will require testing by exercise.</th>
<th>Demonstrate communication and consultation skill with stakeholders. Determine roles, responsibilities and resource implications of involvement exercise.</th>
<th>Achieve and foster commitment from relevant stakeholders (financial and human resources) of involvement in exercise.</th>
</tr>
</thead>
<tbody>
<tr>
<td>Objectives of the exercise which meet the identified need are determined</td>
<td>Determine the objectives of the exercise. Document objectives in clear, simple and measurable terms.</td>
<td>Determine pathways to achieve those objectives.</td>
<td>Determine context evidence required to evaluate stated objectives.</td>
</tr>
<tr>
<td>Exercise style, consistent with the objectives, is selected in consultation with stakeholders</td>
<td>Select the exercise style to meet stated objectives in consultation with stakeholders.</td>
<td>Justify the selection of exercise style to stakeholder groups</td>
<td>Examine the strengths and weaknesses of a range of alternative exercise styles Reviews and modify.</td>
</tr>
<tr>
<td>Exercise design team is assembled</td>
<td>Identify appropriate personnel to design and write exercise.</td>
<td>Assemble and brief exercise writing team and allocate tasks.</td>
<td>Evaluate and provide guidance to meet stated objectives.</td>
</tr>
<tr>
<td>Design exercise</td>
<td>Implement existing exercise formats.</td>
<td>Customise existing exercise formats to suit stated objectives.</td>
<td>Design innovative exercises to meet objectives.</td>
</tr>
<tr>
<td>Resource allocation</td>
<td>Identify required resources.</td>
<td>Justify resource allocation to stakeholder groups.</td>
<td>Secure resources required to implement exercise in consultation with stakeholders.</td>
</tr>
<tr>
<td>Manage exercise</td>
<td>Communicate aims, objectives, expectations and activity outcomes</td>
<td>Initiate and facilitate exercise. Consult with</td>
<td>Monitor and review exercise plan. Provide feedback to</td>
</tr>
</tbody>
</table>
to personnel involved in exercise. partcipating personnel and relevant stakeholders on evaluation of exercise. participating personnel and stakeholders.

Figure 1: The Quality Definition Matrix for the Public Safety Competency Unit

In the 'synthesis project' on multi source competency assessment, we are working with specialists in the public safety (Woods and Cooper, 2002) and public service (Connally and Jorgensen, 2002) industries to identify the differentiating attributes of the candidates. These included strategic and lateral thinking, intellectual demand, autonomy, insightfulness, leadership and intuitiveness. The problem facing the project team was then to develop tasks and recording systems that enabled these differentiating attributes to influence the performance in such a way that clear differences in the quality of the outcome could be observed.

In the assessment and reporting of VET in school outcomes where there is a devolved assessment system of task design and administration. This implies that the professional development needs of teachers / trainers / assessors needs to be considered if changes to assessment design is to be used. Assessors will need assistance in designing differentiating performance tasks, establishing criteria, interpretation and reporting in a standards-referenced framework. The development of differentiating assessment tasks may require skills that extend beyond the skill requirements associated with Certificate IV in Assessment and Workplace Training (NAWTB, 1998). Hence, there would be a need for professional development additional to those units to develop appropriate assessment skills. Possession of the assessment competencies is a necessary, but not sufficient condition for the broader role of assessment in senior secondary schools.

If it is possible to describe levels of performance quality involved in completing assessment tasks, it is also possible to define a range of levels and the criteria that separate them. This, however, requires that there is at least a notion if not a theory of development; knowledge of the kinds of tasks that are indicative of progress; and a capacity to observe different quality of performances on those tasks. These observations occur in circumstances that enable an inference of competence in some specific area beyond the sample of assessment tasks used. These are demanding requirements for assessment developers, but are hardly unreasonable. Central to the task is a linkage of descriptions of performance quality to steps in learning or to components of holistic assessment tasks. An example of a grid used for this design is shown below.

Figure 3 is a grid used in a workshop with those involved in emergency management and is linked to the public safety training package. Across the top shows the broad levels of quality of performance that were first defined using just three levels, low, medium and high. Down the left side is the list of elements that form the unit. The cells of the matrix contain descriptions of the behaviours of the candidates that an assessor would observe in classifying the performance at one of the levels. In this example the element is the ability to use an (emergency management) exercise style that is consistent with the objectives and is selected in consultation with the stakeholders in the community. A candidate operating at a low level in this element would typically select the exercise to meet the stated objectives in consultation with the stakeholders. Another candidate operating at a higher level might also
consult with stakeholders but would also justify the selection. At a still higher level, the candidate might typically examine the strengths and weaknesses of a range of alternatives, review them and modify them where necessary; all the time working in consultation with stakeholders. The assessment would consist of a range of these judgements for each of the elements in this unit.

The public safety project is also examining how these data can be collected from a range of observers and other assessment data, and then synthesising them using an extension of Rasch's (1980) original item response measurement model. The measurement model is referred to as the multi faceted Rasch model (Linacre, 1990) and accounts for each of the systematic influences (such as varying difficulties of the methods and the varying stringency's of the observers), and separates their specific effects from estimates of other influences on measures of candidate's competence. The process of synthesis will control for the effect of different types of assessment, the differentiating levels of stringency characteristic of different judges, and for the differing effects of varied sources of data such direct observation, portfolios, tests, interviews and so on. Each of these is also expected to have systematically different influences on the ratings assigned by different types of observers. The project objective is to provide an assessment of the candidate's quality of performance free of sources of systematic bias.

This procedure has been recognised as having important consequences for the assessment of VET within senior secondary certificates that involve assessments conducted in the workplace, school environment and through central examinations. By controlling for the effects of location, assessor type and for type of assessment, the candidate wishing to be considered for selection in further education or for employment, will not be disadvantaged and can expect to be judged on the basis of the performance, and not the bias introduced by the systematic sources of error inherent in the forms of assessment used.

A second approach to task design is now examined for the hospitality unit, "Perform Clerical Procedures". There are three elements to this unit.

The Unit of Competency: (TTHBFO06A) Perform clerical procedures

Unit Description:

This unit deals with the skills and knowledge required to complete a range of routine office procedures. This unit equates to general skills unit THHGGA02A Perform Clerical Procedures.

Elements and Performance Criteria:

<table>
<thead>
<tr>
<th>TTHBFO06A/01</th>
<th>Process office documents</th>
</tr>
</thead>
<tbody>
<tr>
<td>• Documents are processed in accordance with enterprise procedures within designated timelines.</td>
<td></td>
</tr>
<tr>
<td>• Office equipment is correctly used to process documents.</td>
<td></td>
</tr>
<tr>
<td>• Malfunctions of office equipment are promptly identified and rectified or reported in accordance with enterprise procedures.</td>
<td></td>
</tr>
<tr>
<td>THHBFO06A/02</td>
<td>Draft simple correspondence</td>
</tr>
<tr>
<td>--------------</td>
<td>----------------------------</td>
</tr>
<tr>
<td>• Text is written using clear and concise language.</td>
<td></td>
</tr>
<tr>
<td>• Spelling, punctuation and grammar are correct.</td>
<td></td>
</tr>
<tr>
<td>• Meaning of correspondence is understood by the recipient.</td>
<td></td>
</tr>
<tr>
<td>• Information is checked for accuracy prior to sending.</td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>THHBFO06A/03</th>
<th>Maintain document system</th>
</tr>
</thead>
<tbody>
<tr>
<td>• Documents are filed/stored in accordance with enterprise security procedures.</td>
<td></td>
</tr>
<tr>
<td>• Reference and index systems are modified and updated in accordance with enterprise procedures.</td>
<td></td>
</tr>
</tbody>
</table>

Figure 2: The Office Skills Competency unit

The candidate will receive a set of documents, including an instruction to draft an item of correspondence, process the documents using electronic and manual equipment, store the documents and update the index system.

The assessment task should:

- Be conducted in a work or simulated situation
- Be subject to direct observation
- Be sufficiently flexible to allow industry/office variations
- Allow for adapting from work situations

The candidate needs access to a real or simulated office.

Figure 3: The Holistic Assessment Task for Office Unit

The task designed for this competency was holistic in that it addressed all elements. Hence we make a distinction between holistic tasks and holistic assessment. Holistic tasks can have an analytical scoring process and still lead to an holistic judgement as we shall show with the following examples.

Several steps were involved in the assessment for the candidate. The candidate was asked to receive and examine materials, process them, use the appropriate office equipment and report any problems, deal with variations to the standard procedures, draft correspondence, file materials and adjust the indexing system and finally cope with non standard procedures.

For each of these steps in the assessment, there are recognisable differences in the ways that the candidate performed. For the step dealing with variations for example the candidate might display the following type of behaviour...

1. Variations in documents are not identified
2. Variations identified with assistance
3. Variations are identified independently
4. Variations are reported and assistance requested
5. Variations are noted and modifications are proposed

These show a rise in quality of the performance but do not go beyond the task of noting and adjusting for the variations in documents. However the descriptions given here would allow the assessor to observe distinct levels of quality without having to resort to comparative terms such as "performed at a higher level", "did a better job"

These are called rubrics or rules for recording judgements. Rules for designing rubrics and performance assessments themselves are proposed in another paper (Griffin, 2000). Rubrics provide a means of linking a standards reference framework to dichotomous decisions of competence and to a means of differentiating candidates on the basis of performance quality.

The vertical axis on this grid represents the quality of the performance. It is what we call the quality dimension. The descriptions of differences in performance quality are placed in such a way that reflects their relative quality. Some are placed much higher than others. The anticipated performances for dealing with non-standard procedures are also placed on the grid to reflect their relative performance quality.

We then examine the way the descriptions are placed on the grid and examine the meaning of clusters defined by horizontal slices. The first slice contains a set of descriptions that in fact match most of the performance criteria for the unit and a holistic interpretation of this level is

"Candidate at this level is able to complete all of the standard tasks without assistance and to a moderate degree of accuracy. Follow up procedures are not taken unless they are part of the standard procedures. Variations in document types cannot be handled.

At the next level the holistic interpretation of the descriptions is...

"Follows standard procedures with accuracy and dependability. Non standard documents and procedures can be identified and modifications to procedures handled with assistance and supervision"

The third level is interpreted as

"Accurately completes all standard procedures. Identifies non-standard documents and procedures, and independently modifies approach to handle non-standard items.

These three levels enable a holistic interpretation of the performance quality and one of the levels incorporates all of the performance criteria for the unit.

Scores on each level can be obtained by adding the maximum possible score across columns up to that level, for example for level 2, the candidate can score a 1 or 2 as is the case for five of the columns. The sixth columns only allows a score of 1 up to level 2. Hence the maximum score across columns is 2,2,2,2,2, and 1 or a total of 11. The maximum for level 1 is six so the range for level 2 is a score between 7 and 11. Students scoring between these levels are most likely operating at about level 2 and their performance might be typically described using the overall level description. In a similar fashion, the score ranges for each of the levels are determined as 0-6, 7-11, 12-17 and 18-24. These scores are
modelled on the application of partial credit item response modelling and need to be verified using this process.

Figure 4: The Matrix Linking the Quality Dimension with the Task Components.

This is a more refined approach to that shown for the public safety model in Figure 1. Our task in the public safety project is to develop these grids for a series of competency units and then validate the placement of the descriptions onto the grid using item response modelling. From the validated placements a more accurate holistic interpretation will be available on the quality dimension (Griffin and Gillis, 1999, 2000).

For the office procedure task, the performance criteria can be defined for all of the levels including the lowest level, which does not describe the performance of a candidate who could be considered as competent. What might it look like? It is possible to display all the performance criteria from the 'less than competent' to the top of the continuum. The horizontal line represents the cut point for the competency decision and the higher levels open the score for a differentiating score together with a standards-referencing reporting framework. Notice that grades or ratings are not actually needed. However, this instrument would have to be trialled and validated following the steps outlined above. The relative positions of the indicators (rubrics) and performance indicators within and between levels should be validated empirically using appropriate scaling procedures. This is a quality assurance process that aims to assure the quality of the assessment outcome, but depends on both input and process quality assurance procedures. These quality assurance procedures are discussed later.

Figure 5: Linking Indicators to Levels of Quality

This checklist, slightly extends the current style of commonly used checklists, because it incorporates descriptions of quality in addition to the basic description of the task(s) performed. If the overall components of the competency unit - task performance, task management, contingency management, job/role environment skills and transfer skills - are also included into the performance criteria, clear distinctions will emerge. The validation of these indicators on the continuum can also address the question regarding the developmental or hierarchical nature of these components. This should be another outcome
of the public safety project, but would need to be replicated before any generalisation could be made.

The steps involved might be reduced to three major stages. The definition of the quality dimension is the one that we have concentrated on to illustrate the shift from a dichotomous recording system focusing solely on recognition to a differentiating system that caters for selection but also allows for recognition of competence. The second major stage is the design of a set of holistic tasks that are set in the context of the job role, task management, contingency management and transferability skills. The design of the tasks is important in the development of the standards referenced system and in keeping the integrity of the competency system. The third stage is the implementation of a quality control system that audits the development and management of tasks and procedures, the methods of recording observations and the decisions or inferences that are made based on those records.

Quality Control

Quality control is an area in which the recognition-oriented forms of competency assessment have lagged behind other forms of assessment. There has been a persistent failure to address quality control at the outcome level or at the stage where inferences are made about personal qualities and attributes. Quality Assurance (QA) procedures in a competency based system predominantly address the assessment input and process, rather than the output. For instance, an insistence that the procedures and materials are linked to the assessment guidelines in the training package is entirely appropriate, but is an input control. So too is the approval of assessment plans as is the practice in Victoria for VET in Year 12. The QA shifts to the process and to the assessment outputs when moderation is introduced.

Figure 6: The Taxonomy of Quality Control

The QC process can be described as those linked to input, process and output and this is in turn a hierarchy that improves the quality of the recording and interpreting of the data as we move to a more differentiating process. Input procedures involve the guidelines and approval procedures. Process improvement involves the use of expert judgement, group consensus and the use of exemplars for comparison purposes. Outputs address the assessment...
records directly. These include scaling, calibration and equating of assessment data and assume that the assessment task is an instrument for measuring competence.

Examples

The example is provided for the business training package - Unit BSBCM213A - Produce simple word processed documents. The examples include:

1. Task sample
2. Criteria sample
3. Standards reference band example
4. Examples of the state examination system data analysis and performance bands
5. Examples of how the proposed assessment system would conjointly calibrate with the central examination.

You are working for Travel Australia company as an administration support officer. The company specialises in international travel arrangements. As part of your role, you have to operate varying software applications and provide assistance to other administration staff within your company. You have been asked to word process a letter in response to a customer inquiry on behalf of your supervisor, John Watson. The inquiry relates to a telephone inquiry from Mr. Dixon. The following details were recorded:

![Image of a letter](image)

Your task is to produce a letter in response to Mr. Dixon's inquiry, in accordance with the company's policy -

- Use white A4 paper only
- Set out letters in left alignment, open punctuation style
- At hour turn around
- Use Arial font, 12 point
- Use spell check and grammar check function
- Name and save file using the following conventions for the file address: 'c:\my\document\customer\surname_date' and 'customer\surname_date'

Carefully proof read before printing and submitting work to author for signature.

Figure 7: Sample Practical Task

This example has been developed for the unit in Business services related to the use and development of word processing documents. Two scoring procedures have been requested.
The first takes the quality criteria as a hierarchy. Under this approach the assessor checks just the highest level at which the candidate has been observed. The approach assumes that the quality criteria are ordered categories of performance.

A second approach takes all quality criteria possible to be observed and each one is then checked if observed. For this unit a written task was also developed. Separate scoring rules were developed and these were mapped onto the same standards referenced framework. The quality criteria were placed in a matrix to indicate the relative ability of candidates performing on the job or tasks set for assessment purposes. However, not every task could match every element in a unit of competency. To illustrate this the unit on word processing was assessed using a practical task, a written task and a workplace observation. A portfolio was also designed to assess the candidates against these competencies. Only the portfolio and the workplace observation could cover the entire unit of competence. The other assessment methods had to lead to inferential assessments against other parts of the unit.

<table>
<thead>
<tr>
<th>1</th>
<th>Enters, checks and amends text.</th>
</tr>
</thead>
<tbody>
<tr>
<td>☐ Enters text</td>
<td>1</td>
</tr>
<tr>
<td>☐ Enters text and checks for accuracy of spelling and grammar.</td>
<td>2</td>
</tr>
<tr>
<td>☐ Enters text that is free from error in a document style that enhances appearance and readability</td>
<td>3</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>2</th>
<th>Design and Layout</th>
</tr>
</thead>
<tbody>
<tr>
<td>☐ Uses simple word processing functions to format text and document</td>
<td>1</td>
</tr>
<tr>
<td>☐ Selects from existing styles to format document in accordance with task requirements.</td>
<td>2</td>
</tr>
<tr>
<td>☐ Proofreads documents for accuracy of content and consistency of layout and style.</td>
<td>3</td>
</tr>
<tr>
<td>☐ Creates a new style for formatting document that is consistent with task requirements.</td>
<td>4</td>
</tr>
</tbody>
</table>

Figure 8 Practical Task Scoring Rubrics
John is sitting at his desk. He adjusts his chair up and down and then backwards and forwards, before adjusting the computer screen and keyboard. He then swaps the mouse and pad over to the right hand side of the keyboard and states "Ah...that is much more comfortable". Note that the telephone is on the very edge of the desk, out of reach of John. He does not adjust this. He then begins to type a document that was already placed in a document holder.

As he types, his supervisor has handed him a written letter that must be word processed and couriered within the hour. John, however, thinks to himself about how he will also handle the incoming mail that is part of his routine responsibilities within the organisation, not to mention all his documents that are still in his in-tray awaiting word processing.

1. Describe the steps, in correct order, for adjusting your workstation to suit your ergonomic requirements.

2. Did John adjust his workstation in the correct sequence? Explain.

Steps 3 to 12 follow...

Figure 9: Sample Written Task
1.1 **Workspace, furniture and equipment are adjusted to suit the ergonomic requirements of the user.**

- The student adjusts the workstation...
  - furnishing and equipment to suit personal comfort and needs.

- and describes ergonomic requirements for workspace, furniture and equipment.

- using correct sequencing and cites reasons for ergonomic requirements in terms of OHS consequences.

1.2 **Work organisation meets organisational and OHS requirements for computer operation**

- The student...
  - lists OHS requirements for computer operation

- describes techniques that can be used for implementing the OHS requirements in the workplace

- organises work to abide by OHS requirements as well as workplace requirements

- recommends improvements to own work environment from an OHS perspective to improve safety of oneself and others

---

Figure 10: Sample rating rubrics for workplace observation
Use safe work practices

Confirm document requirements

Produce documents
Figure 11: Match between Tasks and Unit Competency Elements

The Quality Dimension and Task Matrix

The performance criteria are listed across the horizontal axis of the matrix. The relative quality of the performance is listed on the vertical axis. The higher the quality criterion placement the higher the quality of the performance or in other words the greater the amount of ability needed for the performance. It was possible to identify themes within the quality criteria and the themes were described as shown below. Four such themes were identified for this unit of competency and they are presented in the example standards framework bands in the Figure Below.
A distinction can be made between the act of assessment and the act of recording that assessment. It is an important distinction because the record of the assessment is what is used to define and defend the decision about the competence of the candidate. This, in the final assessment of the unit becomes the single set of assessment recording mechanism. The assessor can conduct the assessment using any of the procedures developed as possibilities within this project. Scoring rules associated with the quality criteria have suggested score ranges that could equate to each of the four levels. The assessor need only decide in an holistic manner which of the four levels applies to the candidate and record that level for each unit of competence. The assessment can be based on workplace observation, written assignments, portfolio analysis, interviews or any combination of these. The method of recording is based on this framework and the assessor need only record the level from the reference framework above.

This is the current procedure. The assessor uses a checklist to observe the workplace performance and then records the decision for the unit of competence. It is at present a record of whether the student is not competent or competent or in some instances, a record of a three point scale “not competent, competent or competent with Merit.”
Assessing and Recording - There is a Difference

Performance Bands

The performance levels have definite meaning in terms of the quality of knowledge and skills demonstrated by the candidate. The numbers or codes used to record the levels can always be interpreted in absolute terms. The graded system cannot. This is an important departure from existing systems. The band levels can be treated as part of a stand-alone system. No other information is necessary to identify the competence of an individual candidate. The decision of competence can be made on the basis of the interpretation of the synthesis of underpinning knowledge and skills.

The level codes (a number representing the level assigned) can also be treated as unit scores. If this is the case the unit score can be concurrently calibrated with other assessment information. If the nature of the assessor is recorded as well as the outcome, additional quality controls are possible as pointed out the initial report by Griffin and Gillis (2000). Item response modelling can be adopted to control influences of the type of
assessor and the method of assessment. Where central examinations are used, the unit scores can be concurrently calibrated with item level information.

Item response modelling is not necessary to produce a score but the advantages of its application are immense and strong arguments would be needed to resist its use. For now the simplicity of the assessment recording system is outlined. It is worth stressing again that the assessment procedure can be flexible and the tasks, scoring procedures and quality criteria have been developed to help, not as mandatory procedures. The recording requirement of as assessor would resemble the grid below.

For each unit of competency, the assessor needs only to record the quality level from the standards reference framework. The number of levels may vary from unit to unit. Some in this simulation have 5 level, some have as few as two. Currently all units have two levels (Not Yet Competent/ Competent) or three levels (Not Yet Competent/ Competent /Competent with Merit). This model allows for a variable range of levels and the number of levels depends on the information that can be obtained from the assessment methods. In this example the level code for each unit is recorded and then tallied. Of a possible total of 63, a score of 54 is recorded.

![Figure 13: Teacher requirements](image)
**Linking to Examination Data**

**(i) NSW Examination Analyses**

The NSW Board of Studies supplied examination data for the central examinations for the training packages for four industries- Hospitality, Business Services, Metals and Information Technology. In addition the Board of studies supplied the information regarding the links between the unit element and the examination question. The scoring criteria were also provided. An item response analysis was undertaken at item level and a variable map developed. From this analysis and a skills audit of each item and score point it was possible to produce a direct competency based interpretation of the examination data. From this a series of standards referenced bands could be linked to the exam item interpretation.

This in turn meant that the Interpretation was directly linked to competency units. Six bands were identified using clusters of item responses identified as a result of the analysis.

<table>
<thead>
<tr>
<th>6. Students can copy edit and proof read copy edit documents. They can modify and develop filing procedures for the management of files and records. Students can show how actions demonstrate accountability in preparing financial transactions. Students also have a critical approach to OHS policies and procedures and legislation. These students understand the implications of EEO, industrial relations issues and work conditions.</th>
</tr>
</thead>
<tbody>
<tr>
<td>5. Students can identify and correct errors in text such as spelling, punctuation, grammar and typography and understand the importance of filing procedures in the management of files and records. When using business technology such as word processing packages, students can describe and use the key functions and features. Students also demonstrate an understanding of the importance of accountability and security when processing and preparing financial transactions. Students also have a thorough understanding of OHS policies and procedures, legislation, EEO, industrial relations issues and work conditions specific to the office environment.</td>
</tr>
<tr>
<td>4. Students at this level demonstrate a broad knowledge of the industry, including terminology, relevant legislation, OHS policies, industrial relations issues and working conditions. Students at this level demonstrate the essential skills required for preparing and processing financial/business documents. Students have a basic understanding of the processes involved in records handling, including indexing and filing classifications.</td>
</tr>
<tr>
<td>3. Students at this level communicate using basic industry and enterprise terminology and can identify and interpret standard forms, documents and sheets used in the office environment. They are also able to describe and prioritise typical day to day tasks undertaken in the office environment, and demonstrate a basic understanding of the requirements of working within a team, using business equipment and technology, maintaining information systems and minimising paper wastage. Students have a basic understanding of OHS procedures and legislation specific to the office environment; and an elementary understanding of the procedures required for preparing and processing financial/business documents.</td>
</tr>
</tbody>
</table>
2. Students at this level have an elementary knowledge and understanding of the business office environment, including OHS procedures, legislation, use of business equipment and technology, standard operating office procedures and policies as well as client/customer service expectations. The student has limited use of industry terminology.

1. Students at this level lack knowledge and understanding of the business office environment, including OHS procedures.

These levels result from an interpretation of the skills audit associated with the items and score points in the Variable map presented in the following figure.

Figure 14: Variable Map NSW system

When the proposed model is combined with the NSW examination data when it is analysed in isolation using the level codes as score points, the skills audit associated with every score point is complete. In this figure the candidates are represented by the x on the left side of the figure. On the right of the figure, the codes q20, q17a and so on represent the examination question (q20) and the score point associated with it (q20.1, q20.2 etc.). Unit
11.1, unit 13.1, unit 13.2 etc represent the competency unit and the standards referenced band. The relative vertical location of each candidate indicates the overall ability of the candidate. The relative vertical location of the examination question or the unit score point indicates the relative difficulty of the question or the demand on the candidate to demonstrate the skills involved in the unit score level. The skills audit of the competency unit band is defined by the description of the skills at those levels. Hence it is possible to combine the examination questions with the skills defined in the band level for units and obtain an overall composite description of the skills for candidates at specific levels. This combination allows for a direct interpretation of the knowledge and skills demonstrated by the candidates at school and/or through a central examination. In addition the differentiating score can be derived from the central scale represented by the vertical line. The item response model results in a score that can be directly interpreted into meaningful descriptions of competence through a composite standards referenced framework.

(ii) The Victorian Examination Data Analysis

In Victoria assessors record and report three scores—one for each of three mandatory assessment methods. For each method five generic criteria are used. For each of the criteria, a five point scale enables the assessor to identify and record the quality level for each of the generic criteria. In this example three scores out of 25 are recorded with an overall tally of 62 from a possible 75. The Victorian model can easily be accommodated within the proposed model if the teachers report each of the 15 criteria scores. The difference between the proposed model and the Victorian approach is in the interpretation of the performance and the imposition of the competency decision before a score can be obtained. These issues are discussed in following sections.
In the analysis of the Victorian Examination data, it was possible to obtain a direct interpretation of the multiple choice test item data. However no further analyses were possible since the explicit link between the items and the competency units were not available. Nor was it possible to interpret the data from the generic criteria since the scores were recorded in tallies representing assessment method scores. In order to overcome this an equipercentile equating procedure was adopted to examine the relationship between the school-based assessment and the central examination. This procedure makes the assumption that the two data sets represent measures of the same variable. It should be acknowledged that this might not be the case.

The general alignment of the examination and school based assessments in 2001 and 2002 gives some support to the proposition that students who do well in school based competency assessments, also do well in central examinations. Low scores are aligned in both assessment contexts as well. When this conclusion is reviewed in terms of the skills and knowledge displayed in the central examination there are some concerns that might have to be addressed. The interpretation of the central examination is shown in the table below. Given the parallel natures of the equipercentile distributions it may be that the school-based assessments are indirectly assessing the same knowledge and skills, but there must be some doubt expressed regarding this. An item response analysis would more than likely
illustrate that the skills and knowledge would be interpreted in terms of the general criteria as shown in the variable map below.

![Equipercentile equating of School Based and Central Examinations in Victoria VET](image)

**Victoria- School Based and Exam**

**Figure 16: EQUIPERCENTILE EQUATING OF SCHOOL BASED AND CENTRAL EXAMINATIONS IN VICTORIA VET**

The IRT analysis presented in Figure 17 presents the data in the same format as the NSW analysis. The candidates are represented by the distribution on the left side. The examination questions are represented by the number codes on the right. In a simulation of the 15 scores that could be available from Victorian schools. These are labelled according to the criteria associated with each of the assessment methods available. It emerges that the interpretation would be obtained from a skills audit of the multiple choice items and the generic criteria and scores for each. Some further definition of the criteria and the score points for each would be needed before any meaningful interpretation can be obtained. If the result is a variable map as shown in the figure below, the interpretation will require an audit of combination of exam questions with a skills audit of criteria labelled 'Knowledge', "Organisation", "Ideas", "Expression", "Technology", "Problem Solving", "Planning" and so on. These are important generic criteria and would add valuable information for curriculum planners and employers. If it is possible to combine the two systems a great amount of information as well as a differentiating score would be obtained. The relevant Year 1 examination authorities would be readily able to conduct such an examination and report accordingly with little or no additional resources than are used for the current system.
**Trials**

Trials are to be conducted in 2003 to evaluate the proposed model. It should provide evidence of capability to derive a differentiating score, maintain the competency recognition, check links to curriculum and to competency units. It must demonstrate the capacity to combine multiple sources of assessment data into a single score for university entrance scaling purposes. Moreover the trials must address the procedures required for professional development of assessors and assessment task development so that quality differentiation is possible together with the development of standards referenced frameworks for each unit of competence.

It is essential that as many state agencies as possible are involved in the trials to examine the robustness of the model, and to assess the potential for national collaboration in assessment against national certificates of competence. The trials ought to involve the four participating industries so that resource loads on ITABs and training packages can be estimated.

![Figure 17: Victoria variable map](image-url)
Because of the need to assess the model's capacity to cater for a range of competence, multiple year levels involvement will be needed. This is to ensure that a full range of ability is involved. Given the range of skills exhibited in the central examination scores this might not be a serious concern.

The stakes are high for students involved in the trials. For this reason the data should be accurate, reliable and valid and the trials must provide evidence of this. As a result of this demand large numbers will be needed in the samples used for the trials. A minimum of 400 students would be needed for each participating industry.

At the national working party meeting it was resolved to pilot the standards referenced assessment method in 2002 and to conduct a full trial next year with the following conditions:

1. the trial will cover more states and territories than have been involved in development to this point;
2. the trial will be conducted in workplaces as well as schools;
3. the trial will facilitate more holistic assessment, rather than atomistic competency unit assessment;
4. an evaluation of the impact of VET in Schools assessment on youth at risk;
5. an alternate approach may be developed should the outcomes of the trial indicate that would be preferable.

Conclusion

The end result is a system that allows for differentiation and selection as well as recognition of competence. The central tenet of this approach is the recognition of quality as well as competence. There is nothing to lose in this approach and a lot to gain. Whether we call it grading or some other name such as profiling is a matter of our assessment preference.

References.


Gillis, S, Griffin, P, Trembath, R & Ling, P 1997, The examination of the theoretical underpinning of assessment, A report of a research funded by the Australian National Training Authority research Advisory Council, unpublished University of Melbourne, Melbourne.


Griffin, P. 2000. Performance assessment in a standards based framework. Key note address to the NSW Association of Independent Schools, April 28 Sydney


