Key aspects of competency-based assessment

EDITED BY
WILLIAM C HALL

Foreword by
the Hon Ross Free, MP,
Commonwealth Minister for Schools,
Vocational Education and Training

NCVER

Key aspects of competency-based assessment

EDITED BY
WILLIAM C HALL

Foreword by
the Hon Ross Free, MP,
Commonwealth Minister for Schools,
Vocational Education and Training



DEPARTMENT OF EMPLOYMENT,
EDUCATION AND TRAINING


NCVER
Adelaide 1995

The views expressed in this book are those of the individual authors and do not necessarily represent the views of the Commonwealth Minister for Schools, Vocational Education and Training or the Commonwealth Department of Employment, Education and Training.

© 1995 Commonwealth Department of Employment, Education and Training

ISBN 0 86397335 3
TD/TNC 43.01

Published by
National Centre for Vocational Education Research Ltd
A.C.N. 007 967 311
252 Kensington Road
Leabrook, South Australia 5068

Quality System
Quality Endorsed Company
ISO 9001/AS 3901-1987
QEC Lic. no. 4331
Standards Australia

Printed by Copymaster, Adelaide

Foreword

The foundations for a world-class, competency-based vocational education and training system for Australia are in place, and it is widely agreed that the focus of our effort should now be on implementation.

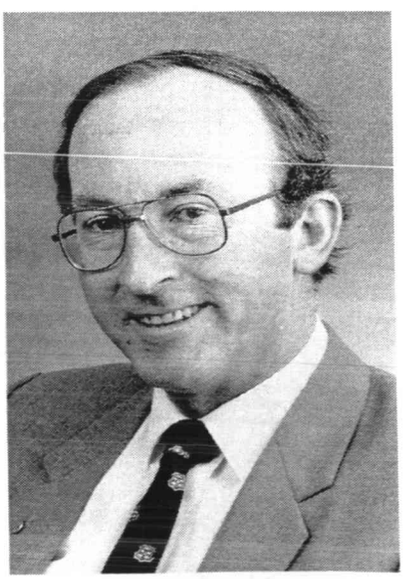
It is also widely agreed that assessment and recognition of prior learning are critical elements of Australia's emerging competency-based vocational education and training system.

Assessment and recognition of prior learning have been subject to considerable debate, yet implementation has lagged behind our expectations, the needs of our industry and the needs of our individuals.

This suite of papers has been commissioned to provide useful, easy to read information for a wide audience within the vocational education and training sector. It is particularly directed to individuals who are addressing the often complex issues associated with implementation of assessment and recognition of prior learning in the workplace.

Each of the papers outlines the findings of research and practice in a topical aspect of assessment. The introductory paper presents a synthesis of the individual papers, and gives us all some food for thought.

I commend these papers to all members of our vocational education and training sector, and hope that they may support substantial implementation of competency-based assessment and recognition of prior learning in the coming years.



ROSS FREE
Minister for Schools,
Vocational Education and Training

September 1995

Contents

Foreword Page

The Hon Ross Free MP, Minister for Schools, Vocational Education and Training iii

Research in competency-based assessment
William Hall 1

Who assesses?
Barbara Bloch, Berwyn Clayton and Jill Favero 9

Validity
Murray Cropley 27

Assessment of knowledge, attitudes and values
Paul Hager 47

Assessment at higher levels
Paul Hager and Shelley Gillis 59

Self-assessment and peer-assessment
Geof Hawke and Patrick Griffin 73

Still counting for nothing?
Bronwyn Travers 85

Needs of special workers
John Saunders and Josie Misko 101

Competency-based assessment in remote locations
Russell Docking 121

Moderation
John Foyster 135

Use of technology
Russell Docking and David Foreman 151

Research in competency-based assessment

William Hall

Introduction

This book of collected papers, commissioned by DEET, follows the popular assessment manuals (also commissioned by DEET) which were published in 1994. In this introductory paper I shall draw upon the other ten titles within the publication in order to make generalisations, so that policy issues may be presented.

Interestingly, we rarely see definitions of 'assessment'. It is assumed that everyone knows what it is and why we do it. Assessment is intended to provide feedback to trainees, to trainers and employers, both during and at the end of a course of training. That feedback shows how successfully trainees have achieved course objectives, nowadays usually expressed as competencies. So, assessment is a way of determining how successfully someone is achieving, or has achieved, (skill) competencies. Put another way: assessment is used to measure competence. The result of that measurement may be used for a variety of purposes, including the awarding of qualifications.

One important statement, by way of background to the whole publication, may be made with conviction: the development of assessment arrangements for competency-based training has been neglected. Much energy has been devoted to the writing of standards and the accompanying curricula. However, without the associated detailed assessment specifications and instruments, competency standards have limited application. This book will be of considerable assistance to those who are responsible for developing assessment procedures and will help to fill an important gap in the competency-based assessment literature.

Also neglected in our development of competency-based training has been the urgent need to train assessors. Bloch, Clayton and Favero deal with this in their paper and I develop some of their thoughts later. Important policy considerations arise from the question: Who should assess? The answer is by no means straightforward and has a bearing on the credibility of assessment procedures.

What we can say (again with conviction) is that all teacher preparation courses, including those for TAFE teachers, should include compulsory major units on assessment and the urgent upgrading (by means of in-service courses) of vocational educators should be a priority. Assessment should be regarded as an important part of the training process and not separate from it.

You will have detected already that in this introductory paper I want to inject a note of urgency! We need to *act*, and to act *quickly*. Therefore DEET is to be commended for its foresight in commissioning the papers which follow, because they cover many of the critical issues.

Subjectivity and objectivity

Essential background to any consideration of competency-based assessment is an understanding of the tensions that inevitably arise between 'objectivity' and 'subjectivity'. Unless this is understood, problems will inevitably follow. The tension is healthy; and it is important to recognise that both subjectivity *and* objectivity are important to competency-based assessment. It is when we fail to recognise this that problems arise.

By 'subjectivity' I mean those processes that require a high element of judgement. They involve people, and the application of their expertise. At the other extreme, we have (for example) machines for marking multiple-choice tests which are wholly objective *at the marking stage.*

This subjectivity/objectivity tension is important and lies at the very heart of many competency-based assessment issues. For example, it means that reliability considerations (generally discussed in terms of objectivity) *do* matter in competency-based assessment; and they cannot be dealt with by assuming that validity (which is dependant on judgement) is the all-important factor. Docking, in his paper, does not fall into this trap: he properly recognises the importance of both validity and reliability, along with flexibility and fairness. Other authors also have this balanced approach, which is essential if the assessment debate is to progress.

Competency-based assessment requirements are discussed in Docking's paper, which quotes from Hager. Hager claims that there are two key features to competency-based assessment: firstly, the requirement to *specify standards*; secondly the requirements to distinguish between *evidence gathering* and *judgement*. Some would try to read into this that judgements (which are subjective) should not be allowed to interfere with the (implied) objectivity of evidence gathering, which is closely linked to specified standards. However, such an assumption ignores the high degree of subjectivity (judgement) involved in first specifying the standards which are to be assessed, including what should (or should not) be specified! My point is that *all* important decisions in life are highly subjective, and this must be recognised even by assessors. In its totality, therefore, competency-based assessment can be just as subjective as (say) setting a traditional academic examination based on essays. Many of the papers recognise this and discuss ways in which the subjectivity of assessment procedures may be minimised.

An important instance of 'front-end' subjectivity is given by Hager in his paper, where he points out that some competency standards are narrowly performance-based, whereas others are too general and not directly linked to outcomes at all. His plea is for holistic (or integrated) assessment and this is supported by other writers.

Only the involvement of competent practitioners (i.e. those who have a depth of knowledge about the skills being assessed, as well as an understanding of assessment) will allow us confidently and professionally to deal with the subjectivity involved in

developing competency standards, as well as the need for objectivity in assessing those standards. Therefore, as a matter of policy, no competency-based assessment framework should be developed or approved without the full involvement of competent practitioners. The standards which are assessed should be developed through a collaborative process followed by a process of validation.

The objectivity-subjectivity debate is at its strongest when considering the question: Who assesses? Bloch, Clayton and Favero have this issue as their topic. The question 'Who?' is inexorably linked to the question 'What for?'

Assessment results may be used to accredit for a qualification a trainee, to provide a licence to practice a trade, and to provide recognition of skills gained so that it is possible to move into a higher award classification. The issue of objectivity versus subjectivity is of critical importance when assessment results are used in these ways. For example, should the assessor be part of the training programme leading to assessment, or separate from it? If the trainer is (even indirectly) paid according to trainee performance (that is, how many trainees are successful) then the pressure to assess the trainee as competent will be substantial or even overwhelming. Objectivity under such circumstances is an unreasonable expectation. So here we have an argument for external assessment, so that the assessor is not part of the training process. On the other hand, if assessment (as distinct from the assessor) is to be part of training, the trainer must be involved in assessing.

The issue of who should assess is a difficult one, and explored by Bloch, Clayton and Favero, who recognise the complex role of assessors, including qualities and qualifications, training and registration of assessors, factors affecting assessors' judgements, and important quality assurance issues. One solution to the problem raised above may be to stop trying to make a single approach to assessment serve a whole range of intentions. The *reason* for assessment should be determined first; the *approach* then follows, including who should assess. Each industry would have to be considered separately, because circumstances may vary from industry to industry.

The most subjectively extreme approach to assessment is self-assessment and peer-assessment, which are described by Hawke and Griffin. The authors are realistic about the limitations of such approaches, also reminding the reader that the techniques of learner self-assessment and peer-assessment are not new. The strength of both techniques possibly lies in the feedback provided to the trainee doing a training course, rather than at the end of a course, when assessment results frequently lead to a qualification.

The reliability/validity debate

All ten papers are concerned with reliability and validity, some explicitly (Docking; Cropley; Hawke and Griffin; Bloch, Clayton and Favero; Foyster; Saunders and Misko) and the rest implicitly. Over many years, numerous books and research papers have been written on validity and reliability; both terms have been dissected, and detailed statistical procedures have been devised. Academic reputations have been made on statements such as 'If you want high validity, you must compromise on reliability'.

Competency-based assessment, because of its very nature, ought to be highly valid. Nevertheless, there is no reason why it should not also be highly reliable. (Subjectivity and objectivity, once again, need not be antipathetical.)

Cropley points out that determining validity is a subjective matter (it relies on expert opinion) which means that competency-based assessment takes on a 'political' dimension. The point is important because there are those who seem to believe that competency-based training and assessment are value-free. They are not (nor should they be). To this end, as Docking and Foreman imply, technology (regarded by some as the ultimate in objectivity) is merely a tool to help support highly subjective policy decisions affecting validity, such as which competencies are important and how general or specific they should be.

Another 'tool', this time consisting of people rather than machines, is the moderator, whose main task is to help improve both the validity and the reliability of assessment. Foyster points out that moderation is a cost-effective, fault-correction device and clearly believes we cannot get too much of it. In his paper he discusses moderation systems in use and realistically evaluates these different approaches.

Hawke and Griffin discuss the difficulties in achieving valid and reliable assessment when using self- and peer-assessment. It seems likely that this assessment approach will be in addition to, not instead of, other approaches. Although not mentioned by the authors, the one thing the general public demand of assessment is 'fairness' and an appearance of high 'objectivity' (spurious though this often may be). Neither self-assessment nor peer-assessment would be regarded by the public as necessarily 'fair' or 'objective' and so, because of this, they have limited application.

It is a healthy sign that reliability and validity are being vigorously debated. As stated already, this objectivity/subjectivity debate is at the heart of competency-based assessment and clearly recognised as such by all of the authors.

Access and equity

Access and equity issues are discussed by Docking, by Travers (who deals with women) and by Saunders and Misko. Competency-based assessment, when applied to workers with special needs and difficulties, has particular requirements. Examples are given by Saunders and Misko, including overcoming language and communication problems, cultural influences and the demands of a changing workplace. A requirement of competency-based assessment is that *all* workers have the right to have their competencies assessed, recorded and recognised. This is a basic difference between traditional assessment and competency-based assessment.

Numerically, as many Australians live in non-metropolitan regions as live in Sydney. Not all of these people live in 'remote' locations (although remote may be defined in different ways for different groups of people). Docking is clear that competency-based assessment offers a unique opportunity for remote learners. Like other authors (but especially Hager) he is a proponent of holistic integrated assessment, but recognises the prohibitive cost of such an approach for remote learners. Therefore, he promotes the assessment of elements of competency, but with evidence of the ability to *link* the separate elements.

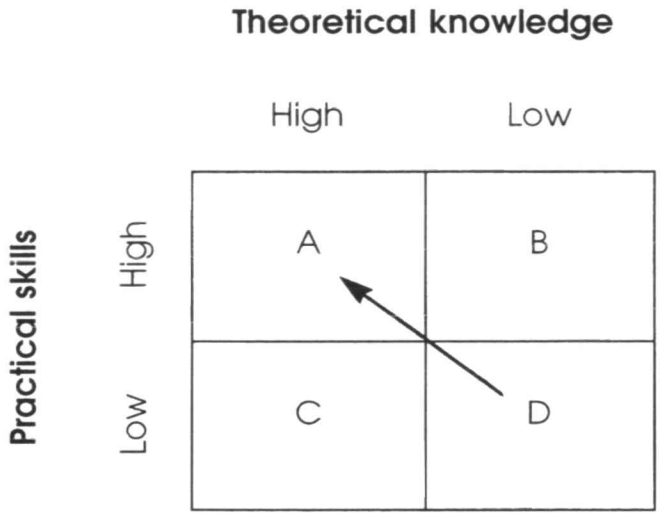
Traditional assessment was used as a hurdle to trip up people: to find out what they did *not* know and, as a consequence, exclude them from the next stage in the learning process. The way to do this was to compare an individual with others, rather than to find out what someone knows (or can do) according to pre-determined standards. And so, in one sense, the old approach was a highly 'political' process. It still is, but in a different sense, because those who have been previously disadvantaged are now able to use competency-based assessment as a lever to gain just a little equity. This is touched upon in Travers' paper.

Travers also states that women's paid employment is often interrupted, a major reason for this interruption being child-rearing. Skill development and its assessment poses special challenges under such circumstances. Of particular concern is the fact that, although women often achieve a high level of competency for a wide range of skills *outside* the workplace, these abilities are rarely recognised *within* the workplace. Those skills are obtained (for example) in voluntary work, as well as in the home. A broader interpretation and implementation of 'recognition of prior learning' would greatly assist in overcoming this problem, a point also acknowledged by Saunders and Misko.

Higher levels of competency

One alleged difference between assessment for competency-based training and assessment for traditional training is the claim by competency critics that only lower level skills may be assessed. The assumption here is that traditional assessment covered higher levels of knowledge, whereas the truth is that mostly it did not (the emphasis was mostly on straightforward knowledge recall). There is no intrinsic reason why high level competencies cannot be assessed: what is needed is the expertise to develop examples of appropriate assessment instruments as exemplars. As a number of authors point out: an holistic approach is essential, which integrates assessment rather than breaking up tasks into small elements, each of which is separately assessed.

Hager and Gillis discuss the assessment of higher level competencies in their paper. *Knowing* that higher level skills may be assessed, and ensuring that they *are* assessed, are two different things. To reinforce this, consider the matrix shown below. There is a considerable danger that in its reform of training, Australia will move from segment D to segment B in the matrix instead of from segment D to segment A. (The matrix draws on OECD documentation.)



The challenge is to ensure that competency-based assessment focusses on segment A, which is one of the arguments in the paper by Hager and Gillis. They point out that higher order competencies are not just confined to the higher levels of the Australian Standards Framework: they are significant at all levels, whenever one is dealing with the less routine aspects of a job.

As Hager states in his paper, an integrated approach to assessment is being used by various Australian industries and occupations in developing their competency standards. This approach is taking account of such attributes as problem solving, communication and analysis. In other words, the competencies are in segment A of the matrix, which is what we should be striving for.

Assessment principles

Cropley quotes the five assessment principles espoused in the National Framework for the Recognition of Training (NFROT):

- competency standards
- flexibility
- validity
- reliability
- recognition of prior learning.

In a strictly statistical sense, these may be incompatible. For example, 'flexibility' and 'reliability' do not usually sit well together statistically. However, the real worry is not mechanical in nature, but is the lack of national action on 'recognition of prior learning'. As a statement of belief, recognition of prior learning (RPL) is now widely held; but faith has yet to be turned into action. Urgent action is needed to convert widely agreed policy and rhetoric into practice.

The five assessment principles are picked up in a variety of ways throughout this publication. The importance of all of them is recognised.

Research

So much 'research' into training is just based on personal opinion or popular rhetoric. That is why this collection of papers is important: they are based on research evidence. Nevertheless, almost all authors show what little research has been done. Their papers recognise that there is a paucity of real, hard research into so much of what matters in competency-based assessment. Here are quotations from five papers that illustrate the point:

> *. . . there is very little research into assessment at remote locations and none at all on competency-based assessment in this context.* (Docking)

> *Further research needs to be conducted to identify examples of best practice in integrated holistic assessment as well as to determine costs that are specific to assessments at higher levels of competence.* (Hager and Gillis)

> *There is, however, no guarantee that any (of these) groups are competent assessors. They may well be, but we cannot be sure because we have not collected any of the evidence to determine this.* (Bloch, Clayton and Favero)

> *Clearly, the question of peer- and self-assessment in competency-based training has not received a great deal of research attention nor has it been used to any degree.* (Hawke and Griffin)

> *We need further qualitative and quantitative analysis of competency-based assessment practices . . .* (Saunders and Misko)

If competency-based assessment is to progress, we need to move on from discussion papers, working party reports, general descriptions and points-of-view. What we desperately need is well-designed and funded *applied research*; but before that is possible we need well-designed competency-based assessment procedures and practices.

The way ahead: Policy findings

All papers which follow in this book spell out the issues and challenges as they affect their topics. Some of these are perfectly straightforward (if costly), such as the need for a national clearinghouse (the current jargon is 'one-stop-shop') to inform and counsel (remote learners) which is promoted by Docking in his paper. Another is the need to train and retrain the trainers so that they themselves are competent competency-based assessors. This is picked up by many of the authors either explicitly or implicitly.

We need lots of examples of best practice; and examples of practical ways to make RPL really work. Although important for all workers, it is especially important for non-mainstream workers.

We need to evaluate more thoroughly the existing competency-based assessment technology (for usefulness and cost-effectiveness) and to provide up-to-date, descriptive and analytical information to assessors about the technology available.

We need to ensure that no national competency standards are approved unless the associated evidence guides are also submitted. Evidence of valid and reliable

assessment instruments, actually in place, should be an essential part of competency standards reaccreditation.

In the same way that we have now separated vocational training *policy* determination from vocational training *delivery*, we need to separate assessment for credentialling and licensing purposes from assessment for trainee feedback. Assessment should be regarded as an essential component of trainee learning (i.e. feedback) as well as a tool for awarding qualifications.

We need to have quality assurance processes in place so that the general public, as well as employers, trainers and trainees, may have confidence in the assessment process.

We need to ensure that we are assessing holistically; which means in an integrated manner, or by ensuring that competency assessment elements are linked and that the ability to make the linkages themselves is assessed. Assessment needs to include attitudes and values as well as knowledge and skills. We need to take a fresh, hard look at the competency standards which have been approved to determine whether they help or hinder holistic assessment.

We need to conduct qualitative and quantitative research into the wide range of assessment techniques. This practical, hard-nosed, research will need industry support, and we will need to convince industry of the benefits.

We need to ensure that competency standards development, curriculum development, assessment and staff development all proceed in an integrated way. Recent national restructuring should help in this: however, assessment must not be left out of the equation.

Above all else, we need to take assessment seriously, realising what an important tool it can be within the training process. Assessment is a critical element of competency-based training and it is encouraging that it is increasingly the focus of our attention. DEET is to be commended for its part in helping with that. However, this is only a beginning: there is a large amount of catching-up to do.

Author

Dr William Hall is the managing director of the National Centre for Vocational Education Research Ltd in Adelaide.

Who assesses?

Barbara Bloch, Berwyn Clayton and Jill Favero

Introduction

The question of who assesses vocational education and training has become increasingly important as industry and the workplace have become more involved in structured training and assessment. In line with the training reform agenda, training providers are developing and running programs that allow for an integration of learning and assessment on- and off-the-job.

This paper identifies and examines issues relevant to assessors in three sectors: industry, TAFE and private providers. Some issues are common to all sectors, while others affect just one training sector. Case studies are used to illustrate many of the issues. Our conclusions point to some future directions and areas needing further research.

The assessment issues identified include:

- the complexity of the assessor role given the different purposes of assessment;
- the qualities and qualifications of assessors in the three sectors
  - assessor standards
  - language and language issues
  - workplace assessor training;
- reliability and validity of assessor judgements;
- quality assurance mechanisms;
- workplace assessment compared with off-the-job assessment;
- small and large employers;
- economic implications.

Definition of an assessor

Links to the purposes of assessment

The Competency Standards Body (CSB) for Assessors and Workplace Trainers stipulates that *assessor* 'is not the title of a position but is the name given to persons acting in an assessment role' (1993, p.v).

In the *workplace*, assessors are likely to be supervisors, leading hands or managers and have many roles to fulfil in addition to their assessment one. Furthermore, assessment in the workplace is often not part of a structured training program.

In contrast, the role of assessors in *training institutions* is to plan and deliver structured courses and, therefore, assessment is an integral part of this. TAFE teachers and community-based providers are also responsible for the development of the individual person, rather than simply meeting industry training needs.

The ultimate purpose of the assessment will also influence who assesses[1]. Purposes can be diverse. For example, assessment may be used to:

- formally recognise an employee's skills which could lead to a higher award classification or wage increment;
- provide licence-related certification;
- provide feedback to students and trainees as part of the teaching and learning process (formative assessment);
- determine whether competency-based learning outcomes of a training program, module or course have been achieved (summative assessment); and
- recognise people's prior learning before entry into a training program, or in order to gain credit for part of it.

(Rumsey 1994, p.33)

Whilst accepting the various purposes of assessment, this paper supports the view that an assessment activity is most effective when integrated into the student or trainee's learning program. That is, assessment should not be seen as an end in itself but should be viewed as an essential part of the learning experience. In this context, formative assessment is particularly important as it provides valuable feedback to the learner and the teacher or trainer. The following case study highlights this aspect of assessment.

Case study 1

Queensland Magnesia (QMAG) provides an example of work-based training using enterprise-specific, self-paced modules developed for the company's production processes. No external training provider was involved and QMAG chose to make a clear distinction between the 'mentor' role and the 'assessor' role.

Mentors were selected for their technical experience and their ability to answer any questions arising from the learning guides. Learners were encouraged to take responsibility for their own learning. Mentors were not necessarily in supervisory roles and had no assessment responsibilities. This was done in order to 'depoliticise' the learning process.

Assessors (usually supervisors in the operational area being assessed) underwent a one-month, in-house, train-the-trainer course. Assessors were expected to be qualified in the theory and practice of the skills and knowledge being assessed, as well as the theory and practice of competency-based assessment.

Complex role of assessors

Making judgements about a person's competence is not simple (Wolf 1995; Guthrie 1993). Even though competency-based assessment may appear 'objective', the reality is that assessors do have to make judgements about when to assess and how much evidence is needed to indicate competence.

Wolf also claims that the more experienced assessors are in their technical field, the more they have 'internalised a model of competence' which may or may not be the same as other people's. As a result, they tend to use 'gut feelings' informed by years of experience as a guide in making assessment judgements. Therefore, it is vital that quality assurance processes are put in place to check that assessors can 'justify their decisions'. Quality assurance is essential for their own protection as well as for the learner and the assessment system as a whole. This issue is discussed further under 'Quality assurance', later in this paper.

Criteria for assessors: Qualities and qualifications

The following list of desirable criteria for selecting assessors is now widely accepted. Assessors should:

- be highly skilled in the aspects of competence they wish to assess;
- have acknowledged competency in assessment and relevant technical areas;
- have an adequate level of literacy in written and spoken English;
- have completed appropriate assessor training;
- be endorsed by the relevant industrial parties.

(Rumsey 1994, p.22)

While the first four criteria are relevant to all assessors, the final one, *be endorsed by the relevant industrial parties,* is essential for workplace assessors. Where assessment is connected to industrial award issues, the recognition of current competencies may be the vehicle for achieving skills recognition and pay increments. In the 1992 research carried out by Broadmeadows College of TAFE on the recognition of prior learning in a number of enterprises 'the sorts of fears documented . . . (on the selection of appropriate assessors) concerned fears about the potential bias of assessors, fears about hidden agendas in the outcomes of the assessments and concerns about the ability of chosen assessors to be fair' (1993, p.6). These concerns may cause some companies to bring in assessors from 'outside' organisations, such as TAFE.

Jenny James, from Tourism Training Australia, spoke at the National Assessment Research Forum (held in Sydney, April 1993) about the hospitality industry's national, industry-wide training system, ACCESS. She stated that 'the selection and training of assessors is perhaps one of the most important components of ACCESS. To become a workplace assessor, ACCESS requires applicants to be highly skilled in the module they wish to assess, possess supervisory experience and have completed a Train the Trainer course'.

TAFE teacher training and ongoing professional development

The criteria identified by Rumsey are equally appropriate for TAFE teachers. Most TAFE systems and colleges require their permanent teaching staff to have vocational qualifications in their discipline area and a specified number of years of industry experience. They are required to have a recognised teaching qualification or be doing a teaching education program. (Interestingly, a recent special TAFE case in Tasmania determined that teaching qualifications there are now only desirable rather than a mandatory requirement.)

A national review of TAFE teacher preparation and development in the late 1980s examined the existing skills and future needs of TAFE teachers. It found that there was dissatisfaction with teacher training courses. Many teachers were also concerned about maintaining their technical expertise and keeping up with the changes occurring in vocational education and training (Hall et al. 1991, pp.15–17).

The 1992 *Staffing TAFE for the 21st century—Phase 1*, found that teacher education courses were not addressing the needs of TAFE teachers and that non-permanent teachers were given little or no preparation for teaching (Predl, pp.15–17). It also identified the need for improved staff development to ensure the long-term maintenance and upgrading of teacher competencies.

In contrast to the qualification prerequisites for permanent staff, the main requirements for the 40 per cent of Australian TAFE teachers who are not permanent, are industrial experience and current technical expertise in their discipline area (Vanderfeen 1992, p.18). Only some of these teachers complete a basic training course before beginning to teach.

Registration of training providers

Under the National Framework for the Recognition of Training (NFROT), any providers of accredited training programs must be registered with their State or Territory Recognition Authority. As part of this process, they are required to give evidence that their training personnel do, at the minimum, meet the competency standards for assessors. Once registered, providers are then responsible for the quality of the assessment they carry out and the maintenance of the assessment records. Registration, therefore, is the main quality assurance mechanism.

How good are assessor standards?

The National Training Board endorsed the Assessor Competency Standards early in 1993. Two years later, the standards were reviewed by the CSB and minor amendments made. The review highlighted that the role of 'assessor' encompasses a number of different functions, not all of which had been incorporated into the original standards. These roles include conducting the assessment (preparing the candidate, carrying out the assessment, providing feedback, recording the results), designing assessment schemes and tools and establishing and managing an assessment system. It was also acknowledged that the skills of designing assessment schemes and managing assessment systems are higher-level skills than those needed for conducting assessments.

Therefore, the CSB has drafted 'Specialist Assessor Competencies', to be undertaken by those people with responsibilities for designing and managing assessments. It is likely that these will only be used by enterprises with a large workforce and a training infrastructure which allows them to have their own assessment system. Smaller workplaces will still rely on 'ready-made' training programs and assessment guides from external sources.

For TAFE, however, assessor training programs based on the revised standards and elective competencies, will be one mechanism for ensuring the TAFE teachers' assessor skills are up-to-date.

Language and literacy

Neither the original standards nor the new specialist standards actually address the language and literacy (and numeracy) implications of workplace assessment. In the industrial context, where assessment may have 'very significant implications for an individual . . . it is especially important . . . that assessors are competent to carry out valid assessments that do not allow any extraneous language and literacy demands in the assessment procedures to stand in the way of the recognition of vocational skills the individual actually possesses.' (Quirk, p.30)

The following is an example of a course designed specifically for learners who may need assistance with English language, literacy and/or numeracy.

Case study 2

NSW TAFE Local Government Foundation Certificate

This is a new certificate and the first for operator-level workers in local government. Its umbrella structure enables it to contain modules from all the areas workers at that level are employed in. Because of the kinds of workers likely to do this training, the curriculum approach was to integrate language, literacy and numeracy into each module. Modules may be assessed on- or off-the-job.

Issues of general interest arising:

- *The assessor may be a TAFE teacher or from the workplace (probably a supervisor), depending on the module being assessed and the specific council (there may be union/employee pressure to bring in an outside provider to do the assessing).*
- *Assessors are required to undertake an endorsed workplace trainers and assessors course.*
- *In rural areas, it is possible that the assessor will come from a neighbouring council.*
- *It is most likely that the assessor will also be the trainer.*
- *The curriculum does not stipulate the proportion of the course to be assessed on- or off-the-job.*
- *If training and assessment for a particular module occurs on the job, then it is the responsibility of both TAFE and the assessor to devise the assessment strategy for that module. It is up to the course co-ordinator and the respective vocational area to provide a quality assurance role.*
- *Assessment verification activities occur by regular meetings and discussions between TAFE and council staff involved; by monitoring the assessment tasks and strategies and the type of evidence gathered; and by 'moderation' meetings between assessors from different councils assessing the same module.*

Is it then the responsibility of assessor-training modules 'to compensate for omissions in assessor competency standards . . . to ensure that assessors are adequately prepared to deal with the language and literacy implications of competency-based assessment?' (Quirk, p.29). Or is this an unrealistic expectation given the brevity of assessor training programs and the fact that, as stated at the beginning of this paper, assessment is likely to be only one of many roles the 'assessor' has to perform during his or her working day?

Workplace assessor training

There is an increasing number of nationally recognised workplace assessor training programs, delivered by private training providers, individual enterprises, TAFEs and universities. They vary in length from two to five days and, at a minimum, train to meet the industry standards.

It could be argued that a few days is not enough time to cover all the components involved in assessment, such as 'ethics, assessor duties and responsibilities, gathering evidence, interpreting and using industry standards, selecting and using forms and methods of assessment' (Rumsey 1994, p.23). Thomson (1994, p.92) queries whether programs of this length can do any more than provide 'survival skills' although acknowledging that employers have difficulties releasing their employees for longer. Thomson also makes the point that we have developed no checks to see if the assessor training programs are, in fact, producing competent assessors.

There is evidence of marked gender imbalance of those attending assessor training. While this is understandable in male-dominated industries such as mining and timber, there is a question about who, in more gender 'balanced' workplaces, is being put forward to attend this training and take on assessment duties[2]. Of equal concern is the apparent lack of assessor trainees from non-English-speaking backgrounds.

The Department of Employment, Education and Training (DEET) currently funds employers up to $500 for each employee who has attended an accredited assessor training program under the Assessor Training Incentive Scheme (ATIS).

Assessors' judgements

Reliability and validity issues on- and off-the-job

As training and assessment is undertaken in many contexts it is difficult for assessors to apply criteria and make interpretations about performance in a completely consistent manner. Reliability of assessment is dependent upon the development of clear guidelines for assessors and general agreement on how criteria are to be interpreted. It is also dependent upon assessors evaluating their assessment and checking to ensure that the assessment outcomes are comparable. An example of such an approach is outlined in the following case study.

Case study 3

A research project is currently underway at the NSW Assessment Centre for Vocational Education comparing two recognition of prior learning (RPL) assessment processes. One RPL process is being performed by an industry assessor and the other by a TAFE teacher. The same group of workers from three oil and chemical companies are being assessed by both assessors.

The RPL processes aim to find out whether applicants are able to be given credit into the Chemical Industry Technology Advanced Certificate by comparing the outcomes of the RPL assessments.

Similarly, TAFE and private providers who are delivering vocational courses totally or partially off-the-job, may have difficulty achieving a high degree of validity in their assessment. For reasons of safety, practicality, cost and simple availability, they often cannot provide learners with extensive real-work experience.

All off-the-job providers of training, therefore, face what Watson (1994) describes as the fundamental problem—having to make judgements about competence on the basis of limited evidence, collected in a setting largely distanced from the reality of the workplace.

In situations where training is being undertaken both on- and off-the-job, the final judgement about competence is best determined in the workplace by workplace assessors. Assessment information relating to skills, knowledge, understanding and attitudes which is gathered away from a real-work setting provides the vital supporting evidence for the final determination of competent work performance (Hager 1994, p.5).

The success of any sharing of assessment is clearly dependent on good communication, trust and a clear understanding of the roles that each of the partners is required to play.

Case study 4

Modules developed by the Joint Australian Public Service Training Council (JAPSTC) for employees of the Australian Public Service were delivered by TAFE (Canberra Institute of Technology) over three days, spaced over a few weeks. In one particular module (not untypical), half the learning took place at TAFE and half through the workplace assessment project took place on the job. The assessor was also the TAFE teacher.

At that time, the workplace supervisor did not have formal assessment duties but was involved in the students' training by suggesting projects for the assessment task, giving advice to the students and also making comments on the assessment cover sheet on the students' work-related project.

Adapted from Bloch and Thomson (1993, pp.93–100)

One other concern about validity centres on the degree to which assessment can effectively predict that a learner is competent and will remain so in future situations. No single assessment can be a perfect predictor of this, but by ensuring that assessment decisions are based on evidence collected from a variety of sources over a period of time, more accurate judgements about competence are likely to be made. In addition, the establishment of verification procedures and studies in this area may help to confirm assessment practices and procedures.

Quality assurance

Verification

The British system of National Vocational Qualifications (NVQs) has a set of quality assurance procedures under the term of 'verification'. There are both internal and external verifiers.

The internal verifier works within an enterprise and has the responsibility of applying quality assurance from within. External verifiers appointed by awarding bodies apply quality assurance across a number of centres (Toop, Gibb and Worsnop 1994, p.124). The verification process includes:

- assessor qualifications and training;
- conducting trials of assessment materials before their widespread use;
- using more than one assessor during the assessment;
- assessors exchanging a selection of evidence and reassessing each others' materials independently, using moderating instruments between assessors. (Toop et al. p.125)

This system has the potential to be bureaucratic and impractical in many business settings. For example, many small workplaces do not have the staff for an 'assessor' role as well as an 'internal verifier' or 'second line assessor' role, or the time to perform all the above quality assurance functions.

In Australia, such formal quality assurance mechanisms have not been developed. Thus, there are questions relating to 'who assesses the assessor?' and 'who is responsible for making the final decision on someone's competence?' The significance of these questions varies according to the context and purpose of assessment. How then can we be assured that assessment at workplace X will produce the same or similar results to workplace Y, when using the same set of standards, or learning outcomes, but with different assessors?

Given that a significant number of teachers have only limited training in assessment, verification is a crucial issue for TAFE. Unlike industry, TAFE colleges should generally have the organisational infrastructure to support some form of internal quality assurance process. Providing skilled assessment co-ordinators and implementing systematic checks of the quality of assessment instruments and the consistency of assessment judgements, would be a good starting point for TAFE institutions.

As with industry, any internal verification process will require resourcing and a clear commitment from teachers and managers to the concept of quality. Effectiveness of verification is further improved when coupled with external verification as outlined in the following case study. The end product may be a raised level of credibility and confidence in TAFE's place in vocational education and training.

Case study 5

Canberra Institute of Technology/Fashion Industry Representatives: Diploma of Arts in Fashion—selected modules.

The curriculum for this course pre-dates full implementation of CBT and there are no industry standards for this sector of the industry. Teachers within the Department of Fashion made the decision to adopt a competency-based approach to the delivery and assessment of the course.

Issues of general interest arising:

- *The assessment strategy was developed by the teachers delivering the course, all of whom have teaching qualifications and extensive experience in industry. Modules were grouped so that an holistic assessment approach could be used. Preliminary discussions took place between the industry parties, the assessment co-ordinator and the teachers delivering the training. The learning outcomes and assessment criteria were discussed and confirmed.*
- *The final assessment decision was determined by a panel of industry representatives drawn from the fashion retail, production and costume design sectors of the local industry. The panel was supported by the TAFE assessment co-ordinator.*
- *For industry, the approach was seen as a meaningful way to have input into the training; for teachers, their expertise and decision-making was confirmed and the students appreciated the instant, useful and relevant feedback.*
- *The strategy was seen to be easy to manage. It was also seen to be valid as it directly addressed commercial fashion development and production in the industry.*

  Additionally, as fashion is an area where judgements about what is innovative or imaginative are easily swayed by the background of the judge, one of the greatest advantages of the approach was that the external people focussed on the outcomes and criteria rather than what was known about the student. As such, the decisions about learner performance were reliable.

  For all participants, the great value of this method of assessment was a high level of confidence in the outcomes of the process.

Work currently being undertaken throughout Australia is looking at models and frameworks to assist workplaces and training providers to work together more effectively than perhaps they have done in the past. For example, the ACTU has developed a Workplace Based Assessment Framework which details the mechanisms by which assessment can occur in the workplace, including responsibility and accountability lines within the workplace and with the registered provider (*The Australian TAFE Teacher*, First Quarter 1995, pp.71–72).

Workplace and off-the-job assessment

Working together

Verification procedures are one aspect of quality assurance. Others include adequate training of assessors and sharing assessment decisions between training providers and industry. This utilises the strengths of both partners in the most cost-effective ways.

Reliability and validity would also be enhanced if the necessary structures were put in place for assessors to exchange 'a selection of evidence and reassess each others' materials independently' (Toop et al. p.125). But, there is a tension between quality assurance and consistency of assessment outcomes on one hand, and devolution and flexibility in training and assessment on the other. The government, through the training reform agenda, is expecting both, although ultimately whether both are achievable across the entire VET system is doubtful.

The following case study effectively demonstrates the use of quality assurance processes and shared decision-making.

Case study 6

BHP, Port Kembla/Illawarra Institute of Technology—Restricted Electrical Licence Course.

This 54-hour module, for non-electrical tradespeople, has traditionally been run off-the-job in a TAFE college. Research was carried out last year by the NSW Assessment Centre for Vocational Education to determine the effectiveness of conducting this course through an integrated delivery of the training and assessment. Two courses were run simultaneously; one taught in the 'traditional' way, the other with an integrated delivery and assessment strategy.

Issues of general interest arising from the 'integrated delivery' group:

- *For reasons specific to BHP at that time, the on-the-job mentors who carried out formative assessment were not the trainees' immediate supervisors. Rather, they were company trainers and supervisors from other areas. In retrospect, it was considered that the mentors should have been closer to the shop floor of the people they were 'mentoring'.*
- *In this pilot, both the TAFE teacher and the workplace assessor were involved in the summative assessments. Although not strictly necessary, and certainly more costly, the exercise increased the trust in the judgements made by both assessors. The company had previous concerns about the standard that TAFE was assessing against.*
- *Because the workers were enrolled in a TAFE course which had to be completed within the semester and were limited by the time constraints of the TAFE teacher, the assessments did not occur as 'naturally' as they otherwise might have done. If assessments can be done within the time constraints of the workplace, worked out between assessor and assessee, the process is likely to be more cost-effective and flexible.*
- *The assessment scheme and instruments were designed collaboratively between TAFE and those involved from BHP. This enhanced the 'joint ownership' and co-operative nature of the assessment process. All assessors (including the TAFE teacher) underwent workplace assessor training.*
- *A few people from the company had concerns about the reliability of the assessments because not all equipment was identical in every respect, the environments in which trainees were being assessed were different and there were different assessors involved. However, it should be noted that all assessors were equally trained and, in particular, the assessment strategies and tools were identical in all cases.*

Quality assurance mechanisms do not necessarily have to be complex and cumbersome in structure to achieve the result of better assessment practice. The following case study highlights some simple methods of verification. It also identifies the learner as an important member of the assessment team.

Case study 7

Maryborough Regional Education and Training Services (MRETS)

Located in central Victoria, this registered, community-based provider runs four accredited courses that are designed to meet the needs of the local labour market. The courses are three Certificates in Occupational Studies—one in Basic Hospitality, another in Business and the third in Office Administration. MRETS also offers a Certificate in General Education for Adults.

Where appropriate, learners are involved in assessing their own performance and this is verified by the trainers. The collaborative approach is seen as a way of motivating learners and encouraging the development of what is a valuable life skill. MRETS trainees are also placed, where possible, into work experience that is relevant to their courses. Employers assist in the assessment of trainees' competencies. They are given clear guidelines about the learning of the outcomes and the assessment criteria. Informal rather than systematic verification of the assessments takes place in the Certificates of Occupational Studies, whilst formal verification occurs with the Certificate in General Education for Adults.

Private providers (particularly community-based providers) are often funded to provide accredited training for participants in labour market programs and assess learning outcomes achieved through off-the-job training rather than through on-the-job experiences. Some programs, however, such as Jobskills or LEAP have a large component of industry experience and can draw on workplace supervisors to assess, or co-assess, participants' competencies.

Small and large employers

One criticism of the training reform agenda is that it is heavily dominated by 'big business' or large employers, and that the training needs of small business have not been catered for. In the 1994 White Paper on Employment, the Federal Government established NETTFORCE to encourage employers to take on people of all ages and pay them a training wage, while providing them with accredited, on-the-job training for a period of 12 months.

A number of States have established a Small Business Industry Training Company (ITC) across a wide range of industry sectors. The role of the ITC is to evaluate whether the employer is 'suitable' to take on a trainee and to develop a training plan for the organisation. Once specific modules and assessment strategies have been identified, the ITC monitors the training conducted at the workplace. Employers are also encouraged to send their trainees to external training, such as TAFE, if that is more appropriate to their needs.

The issue of 'who assesses' in this context is relatively straightforward. The assessor is most likely to be the manager, who may also be the trainer (depending on the

circumstances and size of the business). Assessors are not required to undertake assessor training and many employers consider that they already have the necessary skills and experience to assess their employees' performances. In NSW, the Department of Industrial Relations, Employment, Training and Further Education and the ITC will play a monitoring role in each workplace. Whilst NETTFORCE is yet to fully evolve, the approach is certainly one way of involving small business in the structured training of their employees.

Economic implications

A significant task for managers in all three sectors is to identify ways of making assessment more efficient and cost-effective. There is potential for cost saving if trainers and teachers make more use of holistic assessment tasks. These tasks should provide evidence that trainees have learned what is required in a work situation.

Costs can also be lessened by the use of risk management strategies in the areas of reassessment and the use of less formal supplementary evidence. The challenge is to do these things and still maintain a sense of confidence in the outcomes of the assessment. In part this can be addressed by a firm commitment to the maintenance or enhancement of the technical expertise of teachers and trainers, and providing them with access to training programs which directly target assessment methods and quality assurance procedures.

However, community-based providers have indicated to us that their trainers have not attended workplace assessor training because it has not been offered locally and travel is seen as too costly. This was a factor for the assessors in Maryborough, case study 7. Both industry and training providers must weigh up the advantages of training their personnel to carry out training and assessment, and the value that they will gain from well-assessed skill development.

Future directions

Supporting assessment

Research, resource implications and policy needs

Much of what is written here highlights issues associated with quality assurance and the skills of vocational education and training assessors. Clearly, a considerable amount of assessment is performed by people who are technically skilled themselves but lacking a breadth of understanding of assessment. Equally, many training providers are skilled in assessment techniques, but lack current expertise and experience in their discipline area. Additionally, there is a growing body of assessors who have completed short training courses designed to address the assessor standards.

There is, however, no guarantee that any of these groups are competent assessors. They may well be, but we cannot be sure because we have not collected any of the evidence necessary to determine this. To progress competency-based training further, it is essential that funds are allocated to develop and support quality assessment systems that all participants will respect. Confidence in assessment will only come from a strong commitment to policy development and training based on sound research and ongoing evaluation.

In particular, future research must:

- examine assessor training take-up rates, including identification of mechanisms for enhancing coverage across industries;
- investigate alternative methods for providing assessor training and support both in small business and remote locations;
- evaluate current assessor training programs, including follow-up studies to determine competence of the graduates of the programs and to identify gaps in the training;
- conduct longitudinal studies of the predictive validity of assessment decisions;
- investigate the viability of assessor networks covering the three sectors;
- publish action research undertaken by teachers and trainers to evaluate their assessment approaches;
- disseminate case studies of good practice in integrated training and assessment;
- investigate the practicality and advantages of a national registration of assessors.

Research alone is not going to make a difference. For change to occur there needs to be effective communication of the results to those who will most benefit from the outcomes—teachers, trainers, policy developers and funding authorities. In this way assessors of vocational education and training will be empowered to work together to ensure effective use of resources and so generate change.

Conclusions

When comparing our individual efforts to answer the question of 'who assesses vocational education and training?', the authors found that assessors in industry and in training institutions had many approaches, attitudes and problems in common. While the roles and tasks of assessors may vary, adequate training, maintenance of vocational competence and enhanced skills in assessment are critical issues for all. Additionally, assessors must have a sense that their decisions and processes are supported by adequate structures, training and professional development and a training culture which values assessment.

At the same time, there are clear differences in the roles that assessors play across the three sectors. Most industry assessors have assessment tasks to perform as part of their supervisory duties or they are selected on the basis of their technical expertise. Integrated learning and assessment is basic to the role of teachers and trainers in training institutions.

There is no straightforward answer to the question of who assesses as the purposes, methods and people involved vary according to context. Through the sharing of assessment information, resources and assessment decisions TAFE, private providers and industry may come closer to achieving an integrated, equitable and cost-effective training system. The challenge is to ensure that assessments are carried out by knowledgeable, suitably skilled and experienced assessors. The potential for valid and consistent assessment decisions will then exist.

Endnotes

1 This is not to deny the often stated reality that employers (and managers and supervisors) constantly make informal assessments (judgements) about the competence (or otherwise) of an employee's performance. In this discussion we are referring to those *formal* situations when an assessment occurs at work.

2 Anecdotal information from the approximately two hundred people trained by NSW TAFE in its workplace assessor program suggests that most people are being sent by their organisations so that they can assess for purposes other than for confirmation of course-related learning outcomes. It has also emerged that virtually all the participants were recommended by their supervisors and managers.

References

ANTA Assessment Transition 1995, *Assessment discussion paper (draft)*, March, ANTA, Brisbane.

Australian Training, 1994, 'Investment in Training', vol 1, issue 2, August, pp.12–13.

Bloch, B & Thomson, P 1993, *Working towards best practice*, NCVER, Adelaide.

CSB, Assessors and Workplace Trainers 1993, *Competency standards: Assessors*, PSI consultants.

Davis, S 1993, 'Workplace assessment issues and processes' in *National Assessment Research Forum*, 1/2 April.

Guthrie, H 1993, 'Assessing competence and appraising individuals: Linking the concepts through competency-based training' in *Testing Times*, NCVER/TAFE, Adelaide.

Hager, P 1994, *On-the-job and off-the-job assessment: An issues paper*, Assessment Centre for Vocational Education, St Leonards, NSW.

Hager, P, Athanasou, J and Gonczi, A 1994, *Assessment technical manual*, AGPS, Canberra.

Hall, W C, Dean, S, Guthrie, H, Hayton, G, Mageean, P & Scarfe, J 1991, *National review of TAFE teacher preparation and development—stage 2*, TAFE National Centre for Research and Development Ltd, Adelaide.

Hewett, R, 1995, 'Workplace based assessment framework—the ACTU view', *The Australian TAFE Teacher*, vol 29 no. 1, March, pp.71–72.

James, J 1993. 'Workplace assessment—the ACCESS experience' in *National Assessment Research Forum*, 1/2 April.

Predl, I 1993, *Staffing TAFE for the 21st century—phase 1*, VEETAC Working Party on TAFE Staffing Issues, St Leonards, NSW.

Quirk, R 1994, *Language and literacy issues in competency-based assessment*, Assessment Centre for Vocational Education, Sydney.

Rumsey, D 1994, *Assessment practical guide*, AGPS, Canberra.

Thomson P, 1994 *Current research issues in competency-based assessment*, in National TAFE Senior Executives' Conference, NCVER, Adelaide.

Toop, L, Gibb, J, & Worsnop, P 1994, *Assessment system design*, AGPS, Canberra.

Vanderfeen, M 1992, 'Why competency standards for TAFE teachers?' *The Australian TAFE Teacher*, vol. 26, no. 4, December, pp. 18–25.

Watson, A 1994, 'Three strategies for the assessment of competence', *The Australian TAFE Teacher*, vol 28, no. 1, March, pp. 35–39.

Wolf, A 1995, *Competency-based assessment*, Open University Press, Buckingham.

Authors

Barbara Bloch has a Master in Adult Education (1992) from the University of Technology, Sydney, and is currently a research and planning officer at the Assessment Centre for Vocational Education, TAFE NSW. Her experience in VET research includes quality assurance co-ordinator at the Foundation Studies Training Division, research and development of two outreach courses and work opportunities for women courses and teaching ESOL and literacy for many years, both in Sydney and London. Her publications include *Costing testing*, a paper delivered at the Testing Times (national) conference, November, 1993, Sydney and *Working towards best practice in assessment*, (with P Thomson) 1993, NCVER, Adelaide.

Berwyn Clayton is head of the Curriculum Research and Development Department at the Canberra Institute of Technology. She is a member of both the Australian College of Education and the Australian Institute of Tertiary Education Administrators. Her recent projects have included the evaluation of the Australian Vocational Certificate pilot in Automotive Studies and the development of the curriculum training record books, assessment packages and selection procedures for the Australian Vocational Certificate in Transport and Distribution pilot project. Berwyn's most recent publication is *Focussing on assessment: Strategies for teachers and trainers in vocational education and training*, (NCVER Ltd, 1995) and her *Evaluation kit: Making a world of difference to vocational education and training* will shortly be published by the Canberra Institute of Technology.

Jill Favero, TPTC, Dip Teaching, BEd, is co-ordinator at the Assessment Research Centre, RMIT. She has worked in education and training in a range of fields, from secondary and TAFE colleges to, more recently, industry and labour market program areas. From 1989 to 1994 she was the manager of a vocational education and training centre, clients of which included DEET, State and local government departments, industry and long-term unemployed individuals. Ms Favero is currently co-ordinating a collaborative project with the Comprehensive Adult Student Assessment System (USA). The project involves the development and validation of assessment instruments and the provision of training in key competency areas.

Validity

Murray Cropley

Introduction

Validity is a concept which is very important in all education and training. In vocational education and training (VET) the effectiveness of training programs and competence assessment is directly dependent on the validity (or truth) of the practices followed by the trainers, the trainees and their assessors. In this paper we examine some theoretical issues associated with the concept of validity, and then consider some practical implications of, and ways of appraising, validity. We conclude with a brief look at some possible future directions for improving the validity of competency-based training and assessment.

Validity and related concepts

Validity in an Australian competency-based training context

The National Framework for the Recognition of Training (NFROT) lists five assessment principles which provide the broad guidelines for all assessment in the vocational education and training sector.

Principle 1: Competency-based training

Assessment approaches shall measure the range of knowledge and skills and their application against the competency standards developed by industrial parties and endorsed by the National Training Board, as applicable.

Principle 2: Flexibility in approach

Assessment approaches shall be flexible.

Principle 3: Validity

Assessment approaches used shall be valid, i.e. the techniques used must actually assess what they claim to assess.

Principle 4: Reliability

Assessment approaches shall be able to be relied upon.

Principle 5: Recognition of prior learning

Assessment approaches shall provide for the recognition of competencies held, no matter where they have been acquired. (National Training Board, 1992, p.71)

In this paper we look at the issue of principle 3: validity, but in the context of the importance of the other main principles.

Views about validity

What do we mean by *validity*, particularly as it applies to competency-based assessment? Hager, Athanasou and Gonczi (1994, p.12) state that assessments are *valid* when they assess what they claim to assess.

For example: as a test[1] of ability to count we might hand students a box containing red, blue and green marbles and ask them to count how many of each colour there are. If a person is colour blind, the assessment will depend on colour perception as well as counting ability. In that case, the assessment would not be valid because it did not properly assess what it was claimed to assess.

Messick (1989, p.13) makes the important point that what is to be validated is not the test as such but the inferences derived from test scores. When we talk about the validity of a test we are really talking about the validity of the results got from the test.

Since the early 1950s, test validity has been seen as made up of three main types, with one of these types consisting of two subtypes.

- **Content validity** is evaluated by seeing how well the content of the assessment relates to the area it is meant to be assessing. For example, subject experts make judgements that each question in a test of knowledge and understanding of electric motor design principles is really about design principles and not about repair procedures.

- **Criterion-related validity** is evaluated by comparing the assessment results with one or more external variables (called criteria) considered to provide a direct measure of the area being assessed. The two sub-types of criterion-related validity are *concurrent* and *predictive* validity. Concurrent validity compares two independent assessments. For example, students might do an assignment to assess understanding of the principles of database design by developing a working database. To check concurrent validity, results for this assignment can be compared with results of the same students on a formal assessment on particular aspects of database design. Predictive validity compares assessment results with actual performance. For example, the assessments obtained by trainees during training for a particular job can be compared with ratings of performance by supervisors of those trainees once they have finished the course and are actually doing the job for which they trained.

- **Construct validity** is evaluated by looking at what qualities an assessment measures; that is, by seeing how well certain explanatory concepts or 'constructs' relate to performance in the assessment. For example one of the constructs in an assessment might be seen to be 'problem-solving skills'. For this to be assessed it would be essential that students be faced with a problem which was new to them. If they have prior knowledge of the problem and its solution (e.g. through job experience or having done the same assessment before) the assessment is no longer valid for the construct of 'problem-solving ability'—it assesses knowledge rather than problem solving.

A more recent view of validity combines these three types into one: 'Validity, however, is a unitary concept.' (American Psychological Association 1985, p.8). The single 'type' now favoured is construct validity—validation of a test is validation of its construct(s).

System validity

But what does this mean for us? What is proposed here is that the concept of validity be applicable to the competency-based training *assessment system.* An assessment system is a controlled and ordered procedure designed to carry out the assessment function.

> *The purpose of the assessment system is to provide a means whereby assessment decisions made in relation to many individuals, by many assessors, in many situations can be made consistently, fairly and validly.*
>
> (Toop, Gibb & Worsnop 1994 p.13)

The single type of validity to be considered is still construct validity but it should apply to all aspects of the assessment process and its outcomes. Instead of being worried whether particular assessment outcomes meet some content validity or criterion validity specification, the important thing should be: 'Are the parts of this assessment system valid? Do the outcomes from this system meet our requirements?'

There are essential and optional, although probably desirable, components of an assessment system. Those relevant to validity are as follows:

Essential:

- nominated assessors;
- explicit benchmarks for assessment;
- material to guide assessors in their consideration of evidence;
- record-keeping arrangements.

Desirable but optional:

- licensing, registration and certification of assessors;
- training provisions for assessors;
- verification arrangements;
- appeal mechanisms;
- certification;
- system review mechanisms;
- formal procedures and documentation (Toop, Gibb and Worsnop 1994, p.13).

A valid assessment system should also meet the following criteria:

- assessors must be fully aware of what is to be assessed, as indicated by clearly defined performance or assessment criteria that relate to the units/elements of competency or learning outcomes (which may also incorporate underpinning skills and knowledge);
- evidence must be collected from activities and tasks that can be clearly related to the unit of competency or learning outcomes specified for the course or training program;

- the evidence must show that the performance criteria have been met;
- the evidence must be sufficient, i.e. there has been a suitable sampling of the units/elements of competency or learning outcomes (Hager, Athanasou and Gonczi, 1994, p.12)

These criteria do not, however, cover all that is important when looking at the validity of an assessment system. Indeed, the following principles, based on the NFROT ones, should apply to the four main attributes of competency-based assessment: *validity, reliability, fairness and flexibility.*

These principles are as follows.

Validity principles

1 Assessments will cover the broad range of skills and knowledge needed to demonstrate competency.

2 Assessment of competency should be a process which integrates knowledge and skills with their practical application.

3 During assessment, judgements to determine a learner's competency should, wherever practicable, be made on evidence gathered on a number of occasions and in a variety of contexts or situations.

Reliability principles

4 Assessment practices should be monitored and reviewed to ensure that there is consistency in the interpretation of evidence.

5 Assessors must be competent in terms of the national competency standards for assessors.

Flexibility principles

6 Assessment should cover both the on- and off-the-job components of training.

7 Assessment procedures should provide for the recognition of competencies no matter how, where or when they have been acquired.

8 Assessment procedures should be made accessible to learners so that they can proceed readily from one competency standard to another.

Fairness principles

9 Assessment practices and methods must be equitable to all groups of learners.

10 Assessment procedures and the criteria for judging performance must be made clear to all learners seeking assessment.

11 There should be a participatory approach to assessment. The process of assessment should be jointly developed/agreed between the assessor and the assessee.

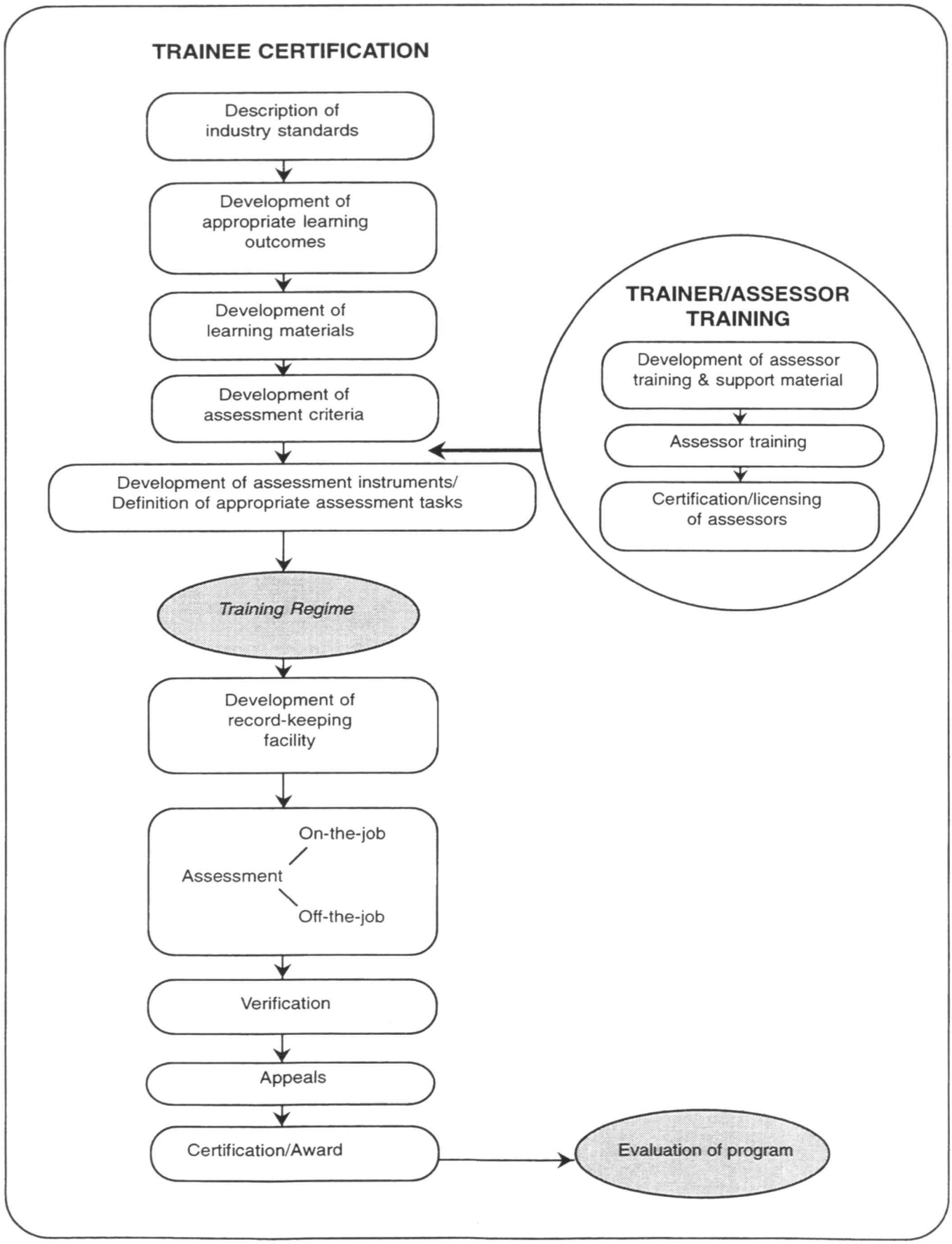
12 Opportunities must be provided to allow learners to challenge assessments and provision must be made for reassessment. (Hager, Athanasou & Gonczi 1994, pp.12–16)

There seems to be a relationship between these principles and most of the components of an assessment system described by Toop, Gibb and Worsnop. For example, the essential component 'explicit benchmarks for assessment' relates to principles 1, 2, 6, 7, 9 and 10.

Therefore it is reasonable to see validity as applicable to a whole assessment system, not just to the assessments arising from that system.

Figure 1 illustrates the major steps of a competency-based assessment system. It is based on the procedures outlined in Toop, Gibb and Worsnop (1994, p.50) and is similar to the model illustrated in *National competency standards* (NTB 1992, p.12). The diagram should be seen as a model—in particular industries some of these steps may not occur, or other main steps may take place.

Figure 1 Major steps required to implement a competency-based training and assessment system



Although it is not a formal part of the certification process, *Evaluation of program* is shown as the last step. (The certification process allows a person to gain a certificate or formal award for their work.) Evaluation is the link between the end of the process and the beginning, and should be a continuous process. It should relate to all steps in the model.

Overall responsibility for making sure that the relevant steps are carried out is shared between industry competency standards bodies (CSBs) and industry training advisory bodies (ITABs). However, particular components of the processes may involve a number of different individuals and groups, such as industry workers, union representatives and TAFE personnel.

Apart from the actual certification/award procedure, all of these steps require their contribution to the assessment system to be validated. We will discuss some of these further in Section 3.

Features of competency-based assessment

There are other concepts which need to be considered in developing a competency-based assessment system and must be addressed as part of the validation process. These are similarly described in Hager, Athanasou and Gonczi (1994), Hall and Saunders (1993), NCCBT (1995), VEETAC (1992) and VEETAC (1993):

- fairness—assessments should not disadvantage particular learners. If an assessment is fair, all learners will be treated equally, will understand clearly what is expected of them and what form the assessment will take. A fair assessment procedure should also support learning—e.g. by providing feedback on assessment performance.
- flexibility—assessments should be flexible enough to cover the range of knowledge and skills covered by the competency standards and to suit the range of delivery modes, sites of delivery and needs of learners. There is no single approach or set of approaches to the assessment of performance in a competency-based assessment system.
- reliability—assessments should use methods and procedures which make sure as far as possible that competency standards and their levels are interpreted and applied consistently to all learners and in a variety of situations. Ideally, a reliable assessment should be reproducible, that is, it should be able to produce the same result for learners of equal competence at different times and/or places.

Other features of a competency-based environment are that assessment should be integrated, holistic and judgemental; that 'the heart of assessment in a competency-based system is the making of judgements of when available evidence is sufficient to satisfy given criteria' (Hager, Athanasou & Gonczi 1994, p.9). Although, by definition, a competency-based system of assessment is based on the concept of competence, we should bear in mind that competence cannot be directly observed. To assess competence we must, instead, base our assessment on observations of task performance.

Performance therefore represents 'the cornerstone of assessment in a competency-based system of assessment' (p.11).

Whilst some researchers believe competence can be judged entirely on performance only, others suggest that there should also be evidence of associated (or 'underpinning')

knowledge: Gibb (1993), for example, states: 'assessment of competence can not rely on one factor alone—performance. Assessment of competence also involves assessment of performance evidence and knowledge evidence combined.' (p.59)

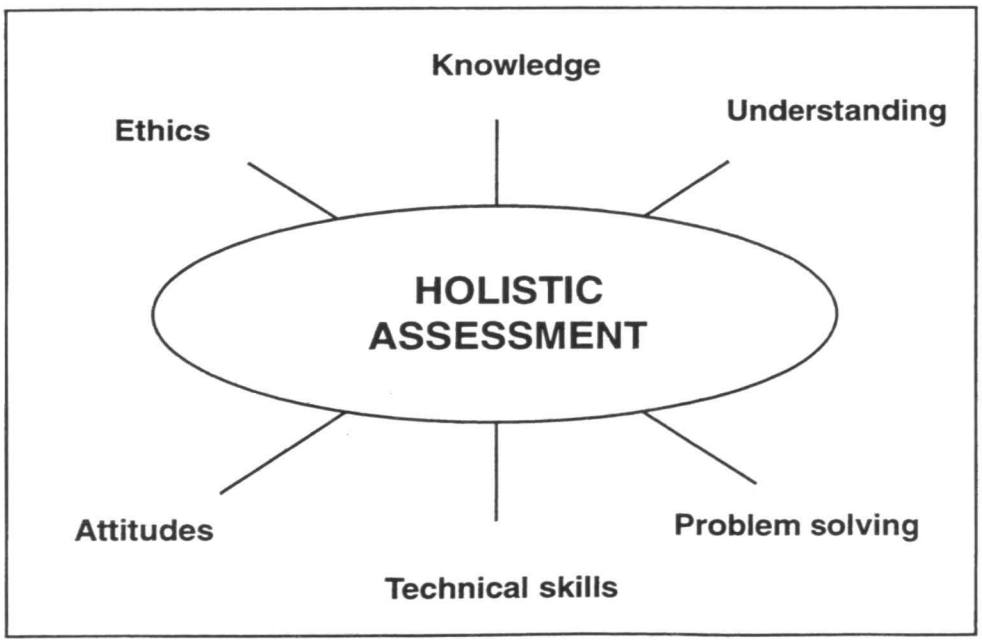
Although other models have been used, nowadays an holistic model of competency is preferred. This means that assessment is based on satisfactory completion of fairly substantial tasks, rather than a series of small ('atomised') tasks or routines. For example, competence in preparing a food course might be determined by noting whether the learner worked safely and efficiently, and produced a satisfactory food dish, rather than whether the right ingredients were chosen, accurately measured, mixed in a specific order, cooked for a prescribed time, and so on.

Hager, Athanasou and Gonczi (1994 p.17) believe that integrated or holistic assessment approaches ought to combine knowledge, understanding, problem solving, technical skills, attitudes and ethics in assessment. Such assessment, they say, can be characterised as:

- problem oriented;
- interdisciplinary;
- embracing actual practice;
- covering groups of competencies;
- focussing on common circumstances;
- demanding analytical abilities; and
- combining theory and practice.

Similar features are highlighted by Gonczi (1993, p.11) in the following illustration:

Figure 2 Unifying features of a holistic assessment model



We should note that not all features need be present in all assessments. For example, assessing competence in using a fork-lift truck may not require assessment of ethics, but it should require consideration of the other five areas.

The other concepts of fairness, flexibility and reliability need consideration in development of a competency-based assessment system and must be addressed as part of the validation process. The principles noted earlier in Section 2.3 provide the basis for this process.

Reliability, in particular, has been viewed as having a special relationship to validity. For example, Hall and Saunders (1993, p.26): 'an assessment must be reliable for it to be valid, but does not need to be valid in order for it to be reliable.' However some believe that the quest for high levels of reliability, by standardising assessment conditions, may lead to decreased validity because the individual workplace environment is part of the real job. Ballenden (1993, p.5) for example, states that trying to ensure that each assessment experience is identical may not mean workplace competence is being assessed.

Nevertheless, reliability as a concept is extremely useful and, where appropriate, assessments should be obtained so that they are as reliable, that is, easy to reproduce, as possible. However, if a choice exists between alternative assessment procedures, which each have at least an acceptable level of reliability, the decision on which to use should be based on the validity of the procedures, rather than their reliability.

Validity in practice

When considering the validity of assessment in practice, we need to examine virtually all aspects of an assessment system, but it is not appropriate to provide here such a large amount of detail. For comprehensive descriptions and examples of a number of assessment systems, see Toop, Gibb and Worsnop (1994). However, some of the steps in Figure 1 are discussed to illustrate particular features of the relationship between an assessment system and its validation.

Validity and description of industry standards

The initial step in defining what skills and knowledge trainees need to demonstrate is the development of industry standards. The National Training Board's publication: *National competency standards* is the most important source of information on this. *National competency standards* provides the fundamental basis for all aspects of the development and use of competency standards. In relation to their development, Chapter 6, 'Content and format of standards', Chapter 7, 'Types of standards', Chapter 8, 'Development of standards' and Chapter 9, 'Standards and avoidance of bias' are particularly relevant. In relation to validity it may be helpful to consider a few of the issues raised there.

In Chapter 6 of *National competency standards* (p.29) four components of competency are identified:

- the requirement to perform individual tasks (task competency);
- the requirement to manage a number of different tasks within the job (task management);

- the requirement to respond to irregularities and breakdowns in routine (contingency management);
- the requirement to deal with the responsibilities and expectations of the work environment, including working with others (job/role environment).

Each of these must be catered for in an industry competency standard, so an important test of the validity of a competency standard would be 'Does it specify for the particular competency the four required components?'

Chapter 7, in relation to the three main types of standards (industry, cross-industry and enterprise), defines industry standards as 'national competency standards endorsed for a specific industry' (p.37). A test of validity of an industry standard would be 'Does this standard apply across a range of enterprises within the one industry, or is it specific to a particular enterprise?'. (It would not matter if the standard also applied across industries; it would certainly matter if it didn't apply to more than one enterprise.)

Chapter 8 poses a number of questions in relation to the actual development of standards.

1 Are the standards able to be readily used at industry and enterprise level?
2 Are the units of competency transferable to a range of contexts?
3 Are the units transferable to a range of jobs?
4 Are the skills most important to the industry identified clearly?
5 Does the standard provide a useable basis for the development of assessment methods and practices, design of courses and recognition of prior learning?
6 Do the standards identify and incorporate relevant competencies involving judgement and understanding?
7 Are the standards forward looking and thus likely to maintain their relevance in the industry?
8 Are the units of competency packaged in a way which facilitates the recognition and portability of skills?
9 Is flexibility in work organisation, competency acquisition and assessment facilitated by the identification, content and packaging of units?
10 Are different levels of competency kept separate through expression in separate units of competency? (NTB 1992, p.42)

'Yes' answers to all applicable questions are needed to make sure that validity issues are being taken care of.

Validity and development of learning materials and development of assessor training and support material

Many training programs involve giving trainees suitable material to help them in their learning. Development of such material so that it relates to the defined standards and prepares trainees for the assessment tasks can be done by professional curriculum developers, in association with industry personnel. It may be most efficient to develop such material at the same time as development of assessor training and support material, so that both trainer and trainee materials conform to common styles and other attributes.

Examples of materials prepared using this approach are the *Workplace trainee workbook* and the *Information for the coach* prepared for the warehousing industry by Western Metropolitan College of TAFE, Victoria. These materials are complementary, as each trainee activity is mirrored by directions for the trainer (called a coach in this material). Descriptions of the required assessment tasks are also given.

The coach's material provides a set of 'coach instructions', which inform the coach of the preparation which is necessary and what instructions the trainee is to be given. The materials also include information about the assessment task for this activity (see Figure 3).

The corresponding trainee material tells the trainee what must be done; provides some helpful tips on how to go about the task; and, for many activities, a sheet where information obtained by the trainee can be entered (see Figure 4). There is also a training record book in which details of successfully completed assessment tasks are maintained.

The following diagrams show sample coach's and trainee's materials for part of the warehousing training program.

Figure 3 Coach's material for a sample warehousing unit

Storage requirements

Coach instruction 3

What do I have to do before I start?

Make a note of any products with special storage requirements and where they are stored.

What do I tell my trainee?

Show the trainee where to find products which have special storage requirements in your workplace, such as:
- goods needing refrigeration in cold store room
- small parts in plastic bins on storage shelves
- dangerous goods in fire safe room
- easily stolen goods in caged area.

What do I tell my trainee to do?

The trainee should complete the 'Storage Requirements' table (Activity 3) in the Trainee Workbook.
This requires:
- listing products with special storage requirements
- locating in the warehouse all the products listed in the table.

When this activity has been completed the trainee can do Trainee Assessment Task 3.

Trainee assessment task 3

Purpose

The trainee is to choose five products which have special storage needs and locate these in the warehouse.

What do I have to do before I start?

Check that the trainee has completed the Storage Requirements' table (Activity 3) in the Trainee Workbook.

Choose five products with special storage requirements. These may be the same as those listed in Activity 3 in the Trainee Workbook.

What do I tell my trainee to do?

Taking each product in turn:
- ask the trainee to tell you the special storage requirements of the product
- show you where to find it

What do I do after the trainee has completed the assessment task?

Tell the trainee if they have successfully or unsuccessfully completed the exercise.

Show the trainee the correct answer, if you need to.

If the trainee has completed the assessment task successfully, put your initials beside **learning outcome number 3 in the Training Record Book.**

Figure 4 Trainee's material for a sample warehousing unit

Storage requirements

Activity 3

What do I have to do?

Fill in the information in the 'Storage Requirements' table on the opposite page, while your coach is showing you products which have special storage needs in the warehouse.
List all the products which have special storage needs.
Write down:
- the name of the product
- special storage needs
- where the product is stored.

This could help

As each workplace will be different, the following list may be helpful.
Write down:
- *bins/binning system*
- *racks*
- *refrigeration/freezers/cold rooms*
- *marked floor space.*

Storage requirements to cater for:
- *size*
- *shape*
- *composition*
- *temperature control*
- *fragility*
- *quantity.*

Storage requirements

Note: an example has been done for you.

Name of product	Storage need	Location
Bag of compost	*Large area/outdoors*	*Behind shed No. 4*

These examples show the development of materials for both trainees and trainers, as well as appropriate assessment criteria, instruments and record-keeping facilities.

Validation procedures would include answering questions such as: 'Does each activity relate to the desired learning outcomes?'; 'Are the materials pitched at the appropriate level?'; 'Do the assessment tasks allow the trainees to demonstrate their knowledge and skills?' For validity, the answer to each question should be 'yes'.

Validity and assessment

There is a lot of information and illustrative material available for validation of on- and off-the-job competency-based assessment procedures. Books such as those by Docking (1991), Hall and Saunders (1993), Hager, Athanasou and Gonczi (1994), and Rumsey (1994); papers such as those by Billam (1991), Linn, Baker and Dunbar (1991), Messick (1994), and Stratton (1991) all provide substantial practical information about how to use validation procedures. Together, those of Rumsey and its companion volume by

Hager, Athanasou and Gonczi provide the most comprehensive and valuable sources of information in the Australian context; the latter gives more detail than the former, but both should be consulted.

Hager, Athanasou and Gonczi provide some very useful discussion on validity and assessment. It provides useful advice on the process of competency-based assessment, including approaches which can be used and their advantages and disadvantages with respect to validity and reliability. The following few paragraphs are taken from Part 1 (pp.12–14).

Validity of assessment is achieved when:

- *assessors are aware of what is to be assessed, as indicated by clearly defined performance or assessment criteria that relate to the units/elements of competency or learning outcomes;*
- *evidence is collected from activities and tasks that can be clearly related to the unit of competency or learning outcomes specified for the course or training program;*
- *evidence demonstrates that the performance criteria have been met;*
- *evidence is sufficient, i.e. there has been a suitable sampling of the units/elements of competency or learning outcomes.*

Validity principles

1 *Assessments will cover the broad range of skills and knowledge needed to demonstrate competency.*
2 *Assessment of competency should be a process which integrates knowledge and skills with their practical application. In general, assessment methods are appropriate to the evidence sought when they closely match the type of performance being assessed.*
3 *During assessment, judgements to determine a learner's competency should, wherever practicable, be made on evidence gathered on a number of occasions and in a variety of contexts or situations.*

Appraising validity

Validity as judgement

How is validity to be appraised or tested? To illustrate the principles, and highlight an important concern, we will look at the example of validation of an assessment.

One of the most common ways of checking assessment validity involves *correlation*: comparing performance, scores or ratings on the assessment task with actual performance on the job. The closer the two sets of results match, the higher the correlation and the higher the validity of the assessment. High validity means the assessment is doing what it was designed to do, and measuring what it was supposed to measure.

But how do we measure actual job performance? This is the crux of the validation procedure: subjective or personal judgements must be made about how well learners perform on the job. Validation of the assessment, therefore, is ultimately based on someone's opinion of how well the actual job is being performed.

The same principles and the same concern also apply to validation of systems of assessment. Here we are interested in determining how well various components of the system are doing what they were meant to do, or how well the system as a whole is doing what it was designed to do. Once again, validation is reduced to a question of subjective judgement—in someone's opinion, the procedure/process/ practice under consideration *does do* what it is meant to do, or the description/definition *is correct*, or the material *is appropriate* for its intended use.

If validation is a matter of judgement, then, the next question is: who makes these judgements and on what grounds?

Implementation issues

The key to successful implementation of any change in behaviour, particularly one on such a vast scale as the introduction across Australia of competency-based training and its assessment, is that the 'hearts and minds' of those responsible for carrying out the work are won. There has been a considerable winning of hearts and minds in the take up of competency-based training.

The fact that these changes *have* been introduced means that the whole movement has been generally accepted. But it has not been without criticism—indeed, it would be hoped that change on this scale would attract some criticism, since it is most unlikely that everything was 'got right' initially. There are concerns about implementation of both a practical and a philosophical nature. An example of the practical is that expressed by Davison (1994, p.36) that the assessment of competency is ideally made 'on-the-job', but there are cost and logistics reasons why this is frequently not possible.

There is no simple solution to this problem; it will probably become less significant with time, however, as more industries and enterprises accept the principles and long-term benefits of CBT, and allow trainees and their on-the-job trainers opportunities which they may not currently enjoy.

Casey (1993, pp.29–30) queries whether on-the-job training and assessment actually guarantees that a particular trainee will be given the experience necessary. He gives the example of a trainee cook who is apprenticed to a pizza restaurant and who would therefore not be able to develop and practise all the requisite skills, compared with a college student who would be exposed to and gain experience in all the defined areas before certification.

This is a case of the appropriateness of the training situation and the validity of the assessment. As with many real-life situations a balance must be struck between following two principles that are desirable, but cannot be attained at the same time—'on-the-job training' and 'coverage of the required material'. A judgement must be made as to which of these two is the more important and the appropriate path then followed. Similar dilemmas will occur in many situations—there is no necessary 'right' answer, and decisions will need to be made on a case-by-case basis as to which of the competing principles (not necessarily these two) is the more important to be supported.

Davis (1993) raises a number of implementation concerns. These are:

- *concerns that assessors might be biased, especially when assessing their peers, or when there are complex personal relationships based at the workplace;*

- *concerns that assessors might enable favouritism or nepotism to occur—especially if there are financial implications of assessments;*
- *concerns that prejudiced views might interfere with the process, ie, that the process must be fair;*
- *concerns that particular groups within the organisation might be disadvantaged by formal assessment processes. Often women or people from non-English-speaking backgrounds are cited here. A commitment to equity is necessary to allay these fears.*

(Davis 1993, p.37)

Griffin (1993, p.70) poses another concern: How do we know that the intentions of curriculum developers in relation to particular training activities are interpreted correctly by the on-the-job trainer and/or assessor?

To return to our earlier question regarding validity appraisal: 'Who makes the judgements and on what grounds?'. The point has been made that validation involves judgement. Making sure that processes *are* valid is therefore a matter of making 'good' judgements. 'Good' judgements are more likely to be made by appropriately qualified people than by unqualified people, and groups of people are more likely to make good decisions than individuals acting alone.

As an example, the accreditation process VET courses go through so that they may be registered under NFROT principles requires groups of people with relevant expertise to determine whether particular courses should be accredited.

Involving a number of suitable people in decision making on contentious issues may provide solutions for the difficulties noted by Davis and Griffin, which were mentioned earlier in this section. For example, Davis's concerns could be addressed if appropriate verification and appeals procedures were established and followed; Griffin's could be addressed by providing support for the on-the-job trainer which might involve other trainers in similar situations, people from TAFE colleges and the curriculum designers.

The grounds for making judgements lie at the heart of the validity appraisal process. Those making the judgements must consider questions of the type noted earlier (3.1) in the discussion of the *National competency standards* Chapter 8. These include questions like:

1 Are the standards able to be readily used at industry and enterprise level?
2 Are the units of competency transferable to a range of contexts?

It is consideration of such questions and the ability to answer 'yes' to them that signifies validity.

Validity and cost-effectiveness

What is at stake

It is likely that a competency-based assessment with high validity will cost more to develop, validate and administer than one with low validity. If we need to assess the competence of an astronaut who is to lead a mission to Mars, the stakes are so high that the expense of very high validity assessments is justified. On the other hand, if we wish to assess the competence of a person who is to deliver leaflets to letter-boxes,

assessments of quite modest validity could be acceptable—a few mis-delivered or damaged leaflets would be of no real concern.

The level of assessment validity therefore is influenced by the purpose of the assessment, or more specifically, the consequences of incorrect assessment. Whilst high validity may be a worthy ideal, in many cases a moderate level of validity at lower cost may be more realistic.

Costs versus benefits—a compromise

Competency-based assessments can be more expensive to administer than the old style of assessment because they need to be administered individually and they can tie up facilities, personnel and equipment for however long the assessment takes. However, properly designed competency-based assessments can also have quite high validity as a measure of a person's ability to perform a job. The decision on the style of assessment, the extent of assessment and the level of validity to accept must be balanced against the cost of assessment and the consequences of incorrectly assessing a person's competence. Cost-effectiveness of assessment therefore involves a compromise between a range of factors. One of these factors is assessment validity.

Industrial implications of validity

Why we need to measure competency

We need to measure competency for a variety of reasons:

- to make sure that a worker's competencies match the job for maximum productivity;
- to maximise workforce flexibility;
- to make career progression easier;
- to enable workers to qualify for extra pay;
- to make sure that workers are able to work safely.

If we accurately assess a worker's competencies, all of these things can happen and the benefits to the worker, the employer and indeed the nation, are maximised. However, if we get it wrong, the consequences can be quite serious.

Consequences of incorrect assessment

Incorrect assessment of competency has the potential to do enormous damage to people, to plant and equipment, to productivity and to worker's attitudes and behaviours.

If we mistakenly judge workers not to be competent it is likely the workers will be dissatisfied; firstly because they are not eligible for career progression and extra pay, and secondly because they feel snubbed by their employer and/or by the assessment system. The company too may suffer through not being able to benefit from full use of the workers' skills, through lowered productivity as a result of the workers' loss of enthusiasm and perhaps through the workers shifting to another employer.

If we mistakenly judge workers to be competent, productivity may suffer because the workers are unable to perform new jobs effectively, accidents may occur causing injury or damage to plant and equipment, and other employees may become dissatisfied and lose respect for the assessment system because they see workers being unfairly promoted.

Many of these outcomes have the potential to cause internal company disputes, involve unions acting on behalf of their members and even the courts in the case of unfair treatment or accidents.

One of the most important keys to preventing these problems lies in the validity of the assessments used. A valid competency-based assessment, properly conducted, maximises the chances of correctly judging a worker's competence to the benefit of all concerned.

Directions and conclusions

Possible future directions

There are certainly areas where improvement in procedures could be introduced into the development of competency-based training and assessment. Some of these might be regarded as efficiency measures. They could also be regarded as affecting the validity of the process.

An example is provided by Griffin (1993). He argues that it may not be necessary to assess all identified competencies for a particular area if these competencies can be placed on a scale of increasing competence. Using an assessment model called 'partial credit analysis' it could be possible to accurately assess trainees on only a sample of assessment tasks.

If assessment procedures could be carried out using a smaller number of assessment tasks than required previously, and the same results were achieved; that is, people were being trained who were equally skilled and knowledgeable as those trained previously, this would mean that the same output was being achieved for less cost. All other things being equal, this would be a desirable result.

Another area where work could be carried out would be in improving the quality of learning materials, both for trainees and trainers/assessors. Since most of the cost of programs involving the use of people is the cost of the people themselves, any way of improving the quality of the work carried out would provide an increase in the validity of the program as well as achieving this at less cost. Significant amounts of work need to be carried out in undertaking the first few steps identified in Figure 1—well-qualified TAFE teachers could certainly assist in these areas. For example, developing appropriate learning outcomes, developing assessment criteria, developing assessment instruments, developing assessor training and support material, and training assessors are all tasks that could benefit from significant input by TAFE teachers with appropriate expertise.

Perhaps the single most significant procedure that could be introduced to ensure or improve validity would be that discussed in the last section: involving as many suitably experienced and qualified people in assessment processes as possible, tempered by considerations of cost and efficiency of decision making.

Conclusion

Validity is a concept that spreads through all aspects of competency-based training and assessment. Procedures which lead to improved validity of any of these aspects should also lead to improvements in the training area and, ultimately, in industry practices. Such improvement would be to the benefit of Australian society and should be encouraged.

Endnotes

1 Although the word 'test' is used, this term is intended to mean any task which is being formally assessed. Similarly, 'test score' does not just mean a score on a formal test, but any assessment, be it a grade, or a number, or categorisations such as 'competent'/'not yet competent'.

References

American Psychological Association, American Educational Research Association and the National Council on Measurement in Education 1985, *Standards for educational and psychological testing*. American Psychological Association, Washington.

Ballenden, C 1993, 'Designing an assessment system for the clerical-administrative workforce', *Testing times conference papers*, National Centre for Vocational Education Research, Adelaide, pp.1–7.

Billam, D 1991, 'BTEC contribution to "assessment and verification of competence"', *Competence & Assessment*, Special edition no. 2, *The Assessment and Verification of Competence*, pp.9–13.

Black, H 1993, 'Sufficiency of evidence', *Competence & Assessment*, 20, pp.3–10.

Casey, D 1993, 'Assessment: The process of quality assurance for education and training', *Testing times conference papers*, National Centre for Vocational Education Research, Adelaide, pp.28–33.

Davis, S 1993, 'Enterprise-based assessment', *Testing times conference papers*, National Centre for Vocational Education Research, Adelaide pp.34–39.

Davison, T 1994, 'Competency-based training (CBT) and Competency-based assessment (CBA): Two criticisms from a reflective practitioner', *The Australian TAFE Teacher*, Fourth Quarter, pp.33–43.

Docking, R 1991, *An A—Z of assessment myths and assessment in the workplace, Competence & assessment*, Briefing Series No. 4, Employment Department, Sheffield.

Gibb, J 1993, 'Trainers/assessors and learners: The human face to the new CBVET system', *Testing times conference papers*, National Centre for Vocational Education Research, Adelaide pp. 57–65.

Gonczi, A 1993, 'Integrated approaches to competency based assessment', *Training Agenda*, 1, 2.

Griffin, P 1993, 'Issues in competency assessment: Don't revisit the seventies', *Testing times conference papers*, National Centre for Vocational Education Research, Adelaide pp.66–82.

Hager, P 1993, 'Principles of competency-based assessment', *Testing times conference papers*, National Centre for Vocational Education Research, Adelaide pp.95–106.

Hager, P, Athanasou, J & Gonczi, A 1994, *Assessment technical manual*, Department of Employment, Education and Training, Canberra.

Hall, W & Saunders, J 1993, *Getting to grips with assessment*, National Centre for Vocational Education Research, Adelaide.

Linn, R, Baker, E & Dunbar, S 1991, 'Complex, performance-based assessment: Expectations and validation criteria', *Educational Researcher*, 20:8; November, pp.15–21.

Messick, S 1989, 'Validity', in *Educational Measurement*, 3rd edition, ed. R L Linn, Macmillan, New York.

Messick, S 1994, 'The interplay of evidence and consequences in the validation of performance assessments', *Educational Researcher*, 23:2; March, pp.13–23.
National Centre for Competency Based Training [NCCBT] 1995, *CBT in vocational education and training*, NCCBT, Richmond.
National Training Board [NTB] 1992, *National competency standards: Policy and guidelines*, second edition, NTB, Canberra.
Rumsey, D 1994, *Assessment practical guide*, Department of Employment, Education and Training, Canberra.
Stratton, N 1991, 'NVQ units of competence and assessment strategies', *Competence & Assessment*, Special edition no. 2, *The Assessment and Verification of Competence*, pp.14–16.
Toop, L, Gibb, J & Worsnop, P 1994, *Assessment system design*, Department of Employment, Education and Training, Canberra.
Vocational Education, Employment and Training Advisory Committee [VEETAC] Working Party on the Implementation of Competency-Based Training 1992, *assessment of performance under competency-based training; administration of competency-based training*, VEETAC Working Party, Darlinghurst.
VEETAC Competency Standards Body—Assessors and Workplace Trainers 1993, *Competency standards assessors*.
Western Metropolitan College of TAFE 1995, *Information for the coach*, Western Metropolitan College of TAFE, Footscray.
Western Metropolitan College of TAFE 1995, *Workplace trainee workbook*, Western Metropolitan College of TAFE, Footscray.

Author

Murray Cropley runs his own consultancy business, HexaDec Services, in Melbourne. He has been a teacher, worked at the Australian Council for Educational Research, the Victorian Institute of Secondary Education/Victorian Curriculum and Assessment Board, been national co-ordinator of the Australian Curriculum Information Network and director of Educational Services at Scanform Computer Systems. Murray's major professional interest has been educational measurement and assessment, and he was significantly involved in the development of assessment procedures for the Victorian Certificate of Education. His current work relates mainly to software design and questionnaire data analysis. Murray is on the executive of the Australian Curriculum Studies Association and edits its newsletter, *Curriculum Perspectives*.

Assessment of knowledge, attitudes and values

Paul Hager

Introduction

There can be no doubt that knowledge plays an important role in competent performance in typical occupations. Competent taxi drivers, for example, not only know alternative routes from A to B, they also know which route is most efficient in peak hour as against off-peak hour, etc. Likewise, competent hairdressers know about a range of hairstyles and appropriate supporting products, thereby enabling them to present a range of options from which their clients can make a choice. However, since competency standards describe successful workplace performance, the focus of competency-based assessment (CBA) is on direct work performance. This provides a sharp contrast with much assessment in traditional vocational education and training courses where the focus is firmly on knowledge.

Like knowledge, attitudes and values also play an important role in competent performance in typical occupations. For example, the values of customer service underpin both competent taxi driving and competent hairdressing. A taxi driver committed to customer service would reflect this in various particular attitudes, for example providing assistance to elderly or disabled persons to help them to be comfortably seated in the taxi cab, assisting customers with their luggage. A hairdresser committed to customer service would reflect this in various particular attitudes, for example taking the trouble to investigate client needs and present options while bearing in mind that many customers are unclear what their real needs are, providing coffee and magazines to customers who are kept waiting. In both cases, the deployment of these attitudes and values represents much more than a commendable concern for others. Rather, these are key factors in bringing clients back for repeat services thereby enabling business to be built up over the longer term.

Competency standards of an integrated (or more inclusive) kind are best placed to represent the complexity of work. Integrated competency standards bring together key work tasks and the attributes needed to perform them. In these kinds of standards, actions and attributes are of equal importance.

In practice, actions and attributes cannot be disassociated without diluting the richness of the competency standards. Attributes include knowledge, values and attitudes, as well as skills and abilities. Equal employment opportunity (EEO) in the workplace represents, amongst other things, a set of values that underpin performance in various jobs. These values in turn imply various attitudes that might be expected to be exhibited in particular situations.

If training and assessing in these values and attitudes are thought about in a non-integrated way, the emphasis is likely to be on a knowledge of principles and values rather than on acting in accord with them. However, when these are thought about from an integrated perspective, the emphasis shifts to the know-how of EEO in actual performance of a given job. Integration results in the values and attitudes figuring in the performance criteria that describe effective performance of the elements of the competency standards. This in turn suggests various training and assessing activities that relate to performance in accord with EEO principles and values as distinct from acquiring information about them.

However, while there is wide agreement that knowledge, as well as attitudes and values of various kinds, are important in the practice of many occupations, there has been in the literature a certain scepticism that these aspects of work performance can be captured in competency standards. For example, Hyland (1993) claims that CBA does not do justice to the range of knowledge required in the quality practice of an occupation and, further, that it grossly misunderstands the nature of knowledge. Specifically he charges that:

> *. . . competence strategies are guilty of two cardinal errors: first, they separate the mental and the physical components of performance, and attempt to appraise these separately; secondly, they mistakenly give performance pride of place in evaluating competence, and seriously underestimate the role of knowledge and understanding.*
>
> (Hyland 1993, p. 63)

Doubt whether values and attitudes can be satisfactorily catered for by CBA has been expressed, for example, by Cairns (1992, p. 15) who asks: 'Is it to be assumed that one can perform a value?'.

It turns out that the Hyland and Cairns' criticisms only apply to bad practice based on a narrow view of competence. If we take a more inclusive, integrated view of competence, not only can knowledge, attitudes and values be captured adequately in competency standards, but CBA offers significant advantages in their assessment. To see this, we need to consider briefly what is meant by a more *inclusive, integrated view of competence.*

Views of competence

Because CBA centres on performance, and since a common view is that performance consists of a series of tasks, CBA is often thought of as assessment of the performance of a series of discrete tasks. Even where work is relatively routine, this 'checklist' approach is dubious since the broader aspects of competent performance, such as planning or reacting to unusual situations, are left out of the picture. Thus, the task view of competence omits higher order competencies from the standards. Hence they are also omitted from any training programs and assessment strategies that are based on these narrow competency standards.

One alternative response is to view competence as possession of a series of desirable attributes, including knowledge of appropriate sorts, skills and abilities such as problem solving, analysis, communication and pattern recognition, and attitudes of appropriate kinds. On this generic approach, training and assessment will be seen in terms of strategies to train and assess candidates in each of these separate attributes.

While this second approach looks more promising as a way of capturing the less predictable variety of non-routine work roles, it also has been widely criticised on the grounds that assessing attributes in isolation from actual work practice bears little relation to future occupational performance. In fact, attributes such as problem solving, analysis and pattern recognition are highly context dependent, so that attempts to teach and assess them out of context are largely misconceived. So training and/or assessing candidates in generic problem solving or communication skills sets up a further problem of how, if at all, candidates will learn to transfer this learning to actual work contexts.

Considerations such as these have led to a different approach, which involves an integrated view of competence. This approach is being employed by various Australian occupations and industries in the development of their competency standards. According to the integrated view, competence is thought of in terms of knowledge, abilities, skills and attitudes displayed in the context of a carefully chosen set of realistic occupational tasks which are of an appropriate level of generality (Gonczi, Hager & Oliver 1990; Heywood, Gonczi & Hager 1992; Ash, Gonczi & Hager 1992; Hager & Gonczi 1993; Hager 1994. See also Biggs 1994).

A notable feature of this integrated approach is that it avoids the problem of myriad tasks by selecting key tasks or elements that are central to the practice of the occupation. The main attributes that are required for the competent performance of these key tasks or elements are then identified. Attributes include cognitive skills (knowledge, critical thinking, problem solving strategies), interpersonal skills, affective attributes and technical/psychomotor skills. Experience has shown that when both attributes and key occupational tasks are integrated to produce competency standards, the results appear to capture the holistic richness of skilled practice in a way that neither of the other two approaches could.

The integrated approach to competence also overcomes the various difficulties posed by Ashworth and Saxton (1990) in their useful catalogue of the limitations of narrow competency standards. (For discussion of this see Hager and Gonczi 1991.) This integrated approach to competence also meets the requirement of the National Training Board that competency standards for occupations go well beyond mere task skills (National Training Board 1992). In short, the holistic and integrated view of competence places attributes in the kinds of contexts in which they are employed in the practice of an occupation. This in turn specifies the kinds of training and assessment situations that are suited to effective learning and assessment.

Knowledge, attitudes and values in integrated competency standards

Having explained the notion of integrated competency standards, we are now in a position to see that the earlier objections posed by Hyland and Cairns in fact provide strong support for the integrated approach to competence. Hyland's first criticism is that 'competence strategies . . . separate the mental and the physical components of performance'. The integrated view of competence asserts just as strongly as does Hyland that any rigid separation of the mental and the physical is misconceived. (For detailed discussion see Hager 1994.)

According to Hyland, separating the mental and the physical results in the assessment of performances that involve mere behaviour and dispositions, thereby leaving no place for 'acts of judgement' which 'presuppose the possession of a wide range of other capacities of a logical, linguistic and conceptual kind.' (Hyland 1993, p. 64) However, it is precisely because the integrated approach insists on holism that such 'acts of judgement' and the like are crucial to successful performance of such everyday workplace activities as, for example, interviewing clients to establish their needs and then presenting them with a series of options from which to choose. Thus Hyland's first criticism is simply irrelevant to typical holistic assessment situations of this kind (see also Leicester 1994).

The appropriate response to Hyland's second criticism, that evaluation of performance seriously underestimates the role of knowledge and understanding, would be that it greatly depends on how holistic is the performance that is being assessed. The trialling of potential new members of string quartets in actual performances is the kind of holistic performance assessment that will normally detect the presence or absence of the required levels of knowledge and understanding. So, the problem lies not with performance assessment as such, but with the kind of performance assessment that assesses discrete physical skills independently, one by one.

Once again, the integrated view of competence agrees with Hyland in rejecting such assessment methods. In fact, as Harre (1979, p. 131) points out, there are two ways in which we can ascribe knowledge to persons on the basis of their performance: we can infer the level of knowledge on the basis of the performance itself and/or on the basis of how the persons account for and justify the performance and related actions. Both ways are important in integrated CBA strategies (see Gonczi, Hager & Athanasou 1993). Given that actual CBA strategies can include traditional knowledge tests as well as holistic performance assessments, there is no basis for any general claim that assessment of competence necessarily undervalues knowledge. So, while no doubt accurately reflecting various *narrow* approaches to competence, Hyland's criticisms in effect *support* the integrated view of competence.

Turning to attitudes and values, Cairns' criticism of narrow CBA also provides indirect support for the integrated view of competence. From the perspective of the integrated approach, Cairns' rhetorical question, 'Is it to be assumed that one can perform a value?', is simply misconceived. This way of posing the problem demonstrates the mindset, so common amongst critics, that competency standards must be inherently atomistic. On this view, adherence to equal employment opportunity principles (say) is yet another box to be ticked. This way of thinking has led some people to assert that competency standards should be restricted to assessment of knowledge and skills. Such a conclusion poses an apparent dilemma for competency standards—omit attitudes and values and be invalid, or include them and be unworkable.

Once again, the integrated approach to competence provides a solution. On this approach, attributes appear in the performance criteria for various aspects of the occupation. There is no major difficulty in describing occupationally relevant attitudes and values in a general way, or in explaining their importance in given workplace contexts.

Experience has shown that where attitudes such as 'empathising with the patient' are important, they can readily be written into the performance criteria for appropriate elements. In fact 'empathising with the patient' is not difficult to assess in real, holistic work contexts where it is an important part of the overall performance. What is

difficult is assessing 'empathy' in the abstract. Thus the integrated approach shares Cairns' scepticism that attitudes and values can usefully be assessed in isolation. However, it turns out that when more holistic aspects of real-work performance are assessed, attitudes and values are often important components of the performance. As will be suggested later, worries expressed by workers' representatives about assessment of attitudes and values reflect fears about possible abuses of assessment of attitudes and values in isolation. The integrated assessment of the holistic approach should serve to allay these fears.

In that competency standards development in recent years has been far from unproblematic, Cairns' and Hyland's worries are not without point. There is solid justification for claims that some standards are overly narrow and instrumental, being solely preoccupied with performance outcomes; while others are too general, lacking any reference to context, thereby being too remote from any actual performance or outcome. The quality of the standards established so far varies, particularly in how well they reflect the complexities of work.

Having established the central roles of knowledge, attitudes and values in plausible competency standards, and, hence, in vocational education and training, it is time to turn to CBA of knowledge, attitudes and values.

Assessment of knowledge in competency-based vocational education and training

Role of knowledge in work

It seems that nearly all work has as its base a body of distinctive knowledge. The amount, type and breadth of knowledge required varies considerably as does the time taken by learners to acquire that knowledge. This variation is due to various factors, including the intrinsic nature of the work itself and the way in which the work is socially organised and divided. A conventional way of describing the labour market regards jobs of higher status or income as those in which workers have both more knowledge and, in particular, more specialised knowledge. In this sense the traditional trades and professions share a common characteristic. The difference between them is less to do with ownership of specialised knowledge and more to do with the mechanisms employed to govern their practice. The trades are generally regulated by externally imposed independent bodies, while the professions are typified by self-regulating structures.

It is important to recognise that knowledge possession, in itself, may not necessarily lead to competent occupational practice. The practitioner has to apply this knowledge appropriately using a range of cognitive abilities (Walker 1992). Even then, the mere possession of the body of specialised knowledge and related cognitive skills does not guarantee that an individual will act competently in work situations. Ultimately it is the application of specialised knowledge and cognitive skills *within practice-based contexts* which leads to competent performance.

It is interesting to note that a recurring criticism of assessment practices in traditional vocational education courses was their undue reliance on the assumption that knowledge was a sufficient indicator of capacity for competent occupational

performance. By assessing learners in situations that closely approximate to actual workplace performance, holistic CBA clearly has the advantage over more traditional methods.

The body of knowledge and its contribution to practice is also problematic on another count. Knowledge is in a continuing state of transformation. For example, much of the knowledge which will be occupying the minds of trainee doctors in the year 2010 has not yet been discovered or codified. Social, technological and legal changes also impact on the knowledge base of practitioners, and these changes are more or less continuous. For these reasons, a number of commentators have suggested that it is more useful to equip practitioners with the skills to locate, evaluate, research and update 'knowledge-in-formation', rather than with the skills required to access and memorise existing knowledge (Boud and Feletti [eds] 1991).

Knowledge is further problematic because of the different ways in which it is created and perceived within occupations. Both theoretical research and practice-based experience produce new knowledge, but the origin of this new knowledge is conveyed not only in the way it is described, but also in the worth it is accorded by members of the occupation. The formal theoretical knowledge domain of occupations comprises various components which include, amongst other things, discipline-based theories and concepts, empirically derived rules and practical principles, and particular propositions about specific cases, decisions and actions. The other knowledge domain is a more personal, practical one involving 'commonsense' or 'tacit knowledge' and is developed through the rigour of everyday practice. This division ensures that there is only limited recognition by the two knowledge producing groups of the contribution made by the other to the practice of the occupation (Oakeshott 1961; Schön 1983; Eraut 1985).

Knowledge and competence

Knowledge and competence are not as disparate as some commentators assume. Wolf (1989, p. 39) argues for the position that competence and education belong together because competency-based education '. . . is perfectly compatible with the learning of higher-level skills, the acquisition of generalisable knowledge (and understanding), and with broad-based courses.' Wolf's reasons for supporting this position include:

- Competence is a construct, and not something that we can observe directly (p. 40), but so too is knowledge. (We infer whether students' knowledge is adequate from their performance on various tests and assignments.)
- What we know of the structure of mind shows the importance of a variety of cognitive abilities. Knowledge recall is only the start. Far from involving practice without theory, as some higher-education critics fear, what competence does is to take us beyond lower cognitive abilities, such as recall, to higher cognitive abilities, such as application and understanding of knowledge.
- Not just that something is done, but why it is done is crucial, that is '"knowing" something involves knowing when to access it, and being able to do so when appropriate—even if it is only in an examination room.' (p. 42).

Holistic CBA does not neglect knowledge precisely because, as Leicester (1994, p. 117) points out, competent performance in real, complex workplace contexts '. . . is competent precisely because it exhibits relevant knowledge and understanding.' Thus, well-planned CBA opens up the possibility of assessing occupational knowledge in more valid ways than has previously been the case.

Assessment of knowledge

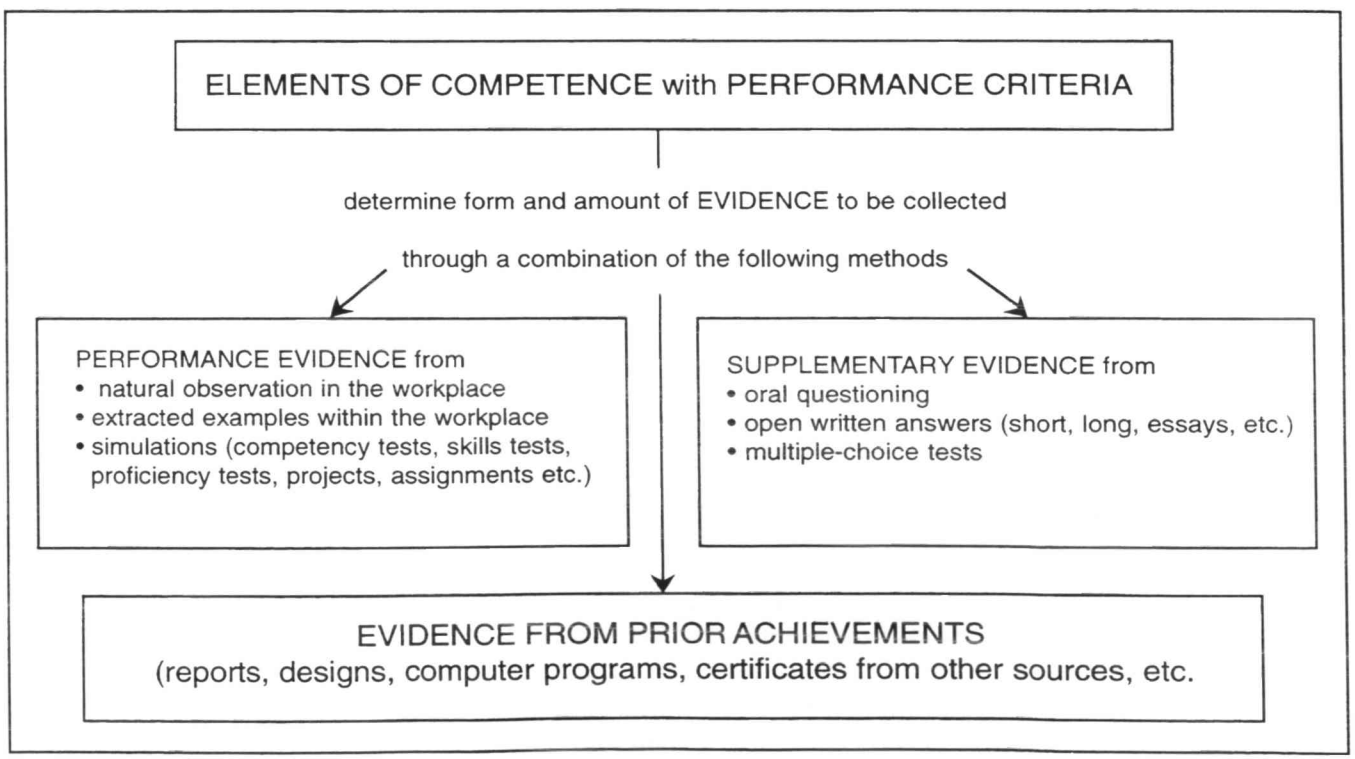
The assessment process is essentially the generation, collection and interpretation of evidence which is then compared to the performance criteria (or standards). This comparison forms the basis of a judgement which infers competent performance or otherwise. The performance criteria (or standards) may be taken as a description of the amount of evidence which needs to be collected to make a 'safe' inference. The ideal is for most of this evidence to be collected from actual work practice. However, owing to such things as cost and time constraints, this will not always be possible in every instance.

Where collection of evidence from actual work practice is not available, the next best sources of evidence are usually those based on simulations of work practice, for example, competency tests, skills tests, proficiency tests. But even these methods may be inadequate to fulfil the range and conditions necessary for adequate assessment of the performance criteria (or standards).

This leaves us with what is sometimes called the assessment 'gap'. The assessment gap is the difference between the amount of evidence which we can reasonably and reliably collect from performance—whether collected from actual work practice or from simulations of work practice—and the amount of evidence which is needed to make a safe inference of competence.

This assessment gap may be filled by supplementary evidence that tests for knowledge and understanding. This additional evidence can be added to the performance evidence to enable a safe inference of competence (see Figure 1).

Figure 1 Overcoming the assessment gap (from Black & Wolf (eds) 1990)



A summary of methods for assessing knowledge in CBA is provided in Hager, Athanasou & Gonczi 1994.

It should be noted that while traditional methods of testing knowledge are so familiar that they have a 'feel' of being inherently more objective than performance assessment, in fact they are used to infer knowledge in much the same way as was pointed out above. Another important point is that a major criticism of traditional methods of testing knowledge has been that they tend to focus on the lower-level aspects such as recall while neglecting higher level aspects like understanding and application. Where knowledge testing is used to provide supplementary evidence for performance assessment these higher-level aspects are less likely to be overlooked.

So, it is possible to argue that there have been major problems with the way knowledge has been traditionally assessed. While the emphasis has been on recall and understanding, other levels of cognition, which include comprehension, application, analysis, synthesis and evaluation, have been relatively neglected. It has never been clear that traditional methods of assessing knowledge actually assess these higher levels of cognition. The problem is highlighted by the fact that people may have fundamental misconceptions about key concepts even after they have passed assessments (Dahlgren 1984). CBA can put candidates into situations in which they are required to comprehend, apply, analyse, synthesise and evaluate data and information. That is, far from downgrading the importance of knowledge, competency-based assessment offers the distinct possibility of better ways of assessing knowledge than have been available through traditional forms of assessment, thereby increasing the validity of occupational knowledge assessment. Clearly, this opens up a fruitful area for informative research into the implementation of CBA.

Assessment of attitudes and values in competency-based vocational education and training

Role of attitudes and values in work

Shared attitudes and values that are regarded as fundamental by practitioners are an important distinguishing feature of many occupations. They are passed on to novices through the formal and informal learning practices which characterise the initiation of new members into the occupation. Teachers, for example, believe that the capacity to learn and understand is a fundamental right of people. This engenders attitudes about such matters as admission, assessment, etc. Likewise, many medical occupations believe that health is a fundamental human right. Residential care workers working with people with developmental disabilities regard respect for the individual as a paramount value within the occupation. Financial services employees regard honesty and trust as important practice-based values.

These values, though different for different occupations, all require implementation and application to the day-to-day practice of work within the occupation, and it is in the area of real practice that values are displayed and reflected in the attitudes of workers in particular situations. It is now possible to make judgements concerning how successful occupations have been in reflecting their values base in the standards that they have developed. Occupations which are vocational in the traditional sense, and

which are active in delivering a range of human services to society, are those which are likely to have particularly strong values bases. Chappell & Hager (1994) provide examples of the role of values and attitudes in the work of youth sector workers and residential care assistants. They also describe how these values and attitudes figure in the competency standards.

If competency standards' development is participative, it has the added advantage of providing practitioners with an opportunity to challenge, clarify and legitimise existing, generalised, taken-for-granted occupational values and to translate them into agreed and acceptable contextualised practice.

Experience so far in competency standards development suggests that there is suspicion of and strong resistance to assessment of discrete, isolated attitudes and values. However, such suspicion and resistance dissipates once core attitudes and values are assessed as underlying components of performance in holistic assessment situations. So, for example, unions are rightly unimpressed at the prospect of a worker being deemed incompetent due to 'having the wrong attitudes to safety'. However the case of someone performing in an unsafe manner when carrying out real work activities is an altogether different matter.

Assessment of attitudes and values

The integrated approach to competence is the key to assessing attitudes and values in a way that is non-threatening to workers. This can then avoid spurious industrial disputes founded on suspicions about arbitrary and idiosyncratic 'hidden assessments' of individuals' attitudes in isolation from job performance. On this approach, attributes appear in the performance criteria for various aspects of the occupation. There is no major difficulty in describing occupationally relevant attitudes and values in a general way, or in explaining their importance in given workplace contexts.

As noted earlier, experience has shown that where attitudes such as 'empathising with the patient' are important, they can readily be written into the performance criteria for appropriate elements. In fact 'empathising with the patient' is not difficult to describe and assess in real, holistic work contexts where it is an important part of the overall performance. What is difficult is describing and assessing 'empathy' *in the abstract*.

None of the foregoing denies that there are real difficulties for researchers who attempt to produce measures of attribute acquisition by individuals. Attributes such as empathy, patience, creativity, honesty, respect and compassion do not lend themselves to simple measurement. Thus attempts to measure values and attitudes by, for example, standard pen-and-paper tests, are notoriously unreliable and of dubious validity. Rather, valid assessment of attitudinal and value factors requires longitudinal and multiple assessments that gather evidence of attitudes and values from a variety of sources. Such sources include the following:

- direct observation of work activities
- supervisor assessments or ratings
- evidence from prior achievements
- oral questioning
- written tests

- self-reports
- practicum, internship, professional year.

These assessment suggestions accord with a fundamental principle of CBA that integrated or holistic assessments should be used wherever possible to cover multiple elements and/or units from the competency standards. This enables the assessment to be more valid, whilst at the same time saving on costs and being more efficient.

The first step in choosing from the available assessment methods is to ascertain which ones are most capable of assessing competence in an integrated manner. Integrated approaches seek to combine knowledge, understanding, problem solving, technical skills, attitudes, and ethics in assessment. Integrated or holistic assessment can be characterised as:

- problem orientated
- interdisciplinary
- embracing actual practice
- covering groups of competencies
- focussing on common circumstances
- demanding analytical abilities
- combining theory and practice.

Far from attitudes and values highlighting a weakness of competency standards development, well-conceptualised standards will turn out to capture attitudes and values in real workplace contexts more effectively than any of the alternatives. Hence, competency standards will enable us to improve upon previous attempts to assess attitudes and values in vocational education and training programs. Once again this points to a promising area for further research into implementation of CBA.

Costs of competency-based assessment

An important practical aspect of the implementation of CBA is the cost. How costly is CBA? How does its cost compare with the cost of the alternatives?

The first point that needs to be made is that any given example of a CBA system, in common with all assessment systems, will represent a compromise. In all assessment systems such factors as validity and reliability can always be further enhanced by the effective deployment of more time and resources. In practice, assessment systems are designed so that a balance is struck in which acceptable levels of validity, reliability, etc. are achieved at an affordable cost.

So, in the implementation of CBA choices on matters such as how to gather the assessment evidence will have cost implications. For example, element-by-element assessment will generally be more costly than more holistic forms of assessment that integrate elements. Fortunately, the more holistic forms of assessment are generally more valid anyway. So, in planning a CBA we can go for a 'deluxe' system at one extreme or a basic 'no-frills' system at the other. In most cases, the aim will be to keep costs within reasonable limits whilst achieving to a satisfactory degree the benefits of CBA.

Interviews with curriculum developers and vocational teachers suggest that a course can become competency-based yet be no more costly in its assessment strategies than the more traditional course that it is replacing. Likewise, in a series of recent case studies of CBA diverse examples representing a range of different contexts were judged to be both effective and of modest cost (Spectrum Research and Training 1994). This has also been the experience with CBA for accreditation of specialist solicitors in NSW, a system which is now in its third year of operation. Cost effective ways of implementing CBA seem likely to be promising research areas in the next few years.

There is currently no reason to suggest that CBA is any more expensive than the alternatives. The initial cost of setting up a CBA system should be offset against the significant benefits, including better learning, that CBA offers over its rivals. Once a CBA system is in place, its ongoing cost should be minimal. (All assessment systems involve ongoing costs. For example, costs of running written examinations include the setting of the examination, having it printed, hiring examination rooms, supervision of the examination, and marking of the examination.)

Conclusion

This paper has argued that an holistic, integrated approach to CBA enables the valid assessment of the knowledge, attitudes and values that play an important role in competent performance in typical occupations. The holistic, integrated approach to CBA has been explained and the principles for applying it to the assessment of knowledge, attitudes and values have been outlined. It has been suggested that good practice in holistic, integrated assessment of knowledge, attitudes and values is a promising area for further research into the implementation of CBA. Finally, it has been argued that the costs of implementing this kind of CBA are comparable to alternative types of assessment.

Bibliography

Ash, S, Gonczi, A & Hager, P 1992, *Combining research methodologies to develop competency-based standards for dietitians: A case study for the professions*. Research Paper No. 6, National Office of Overseas Skills Recognition, DEET, Australian Government Publishing Service, Canberra.

Ashworth, P D & Saxton, J 1990, 'On *competence*', *Journal of Further and Higher Education*, vol. 14, pp. 3–25.

Biggs, J 1994, 'Learning outcomes: Competence or expertise?', *Australian and New Zealand Journal of Vocational Education Research*, vol. 2, no. 1, pp. 1–18.

Black, H & Wolf, A (eds) 1990, *Knowledge and competence: Current issues in training and education*, Scottish Council for Research in Education, Edinburgh.

Boud, D & Feletti, G (eds) 1991, *The challenge of problem based learning*, Kogan Page, London.

Cairns, L 1992, 'Competency-based education: Nostradamus's Nostrum', *Journal of Teaching Practice*, vol. 12, no. 1, pp. 1–32.

Chappell, C & Hager, P 1994, 'Values and competency standards', *Journal of Further and Higher Education*, vol. 18, no. 3, pp. 12–23.

Dahlgren, A 1984, 'Outcomes of learning' in Marton, F, Hounsell, D & Entwistle, N (eds) *The experience of learning*, Scottish Academic Press, Edinburgh.

Eraut, M 1985, 'Knowledge creation and knowledge use in professional contexts', *Studies in Higher Education*, vol. 10, no. 2, pp. 117–133.
Field, J 1991, 'Competency and the pedagogy of labour', *Studies in the Education of Adults*, vol. 23, no. 1, pp. 41–52.
Gonczi, A, Hager, P & Oliver, L 1990, *Establishing competency-based standards in the professions*. Research Paper No. 1, National Office of Overseas Skills Recognition, DEET, Australian Government Publishing Service, Canberra.
Gonczi, A, Hager, P & Athanasou, J 1993, *The development of competency-based assessment strategies for the professions*. Research paper no. 8. National Office of Overseas Skills Recognition Research, DEET, Australian Government Publishing Service, Canberra.
Hager, P 1994, 'Is there a cogent philosophical argument against competency standards?', *Australian Journal of Education*, vol. 38, no. 1, pp. 3–18.
Hager, P & Gonczi, A 1993, 'Attributes and competence', *Australian and New Zealand Journal of Vocational Education Research*, vol. 1, no. 1, pp. 36–45.
Hager, P, Gonczi, A & Athanasou, J 1994, 'General issues about assessment of competence', *Assessment & Evaluation in Higher Education*, vol. 19, no. 1, pp. 3–16.
Hager, P, Athanasou J & Gonczi A 1994, *Assessment technical manual*, DEET, Australian Government Publishing Service, Canberra.
Hall, W C 1994, *Competency-based training and assessment*, National Centre for Vocational Education Research, Adelaide.
Harre, R 1979, *Social being: A theory for social psychology*, Basil Blackwell, Oxford.
Hyland, T 1993, 'Competence, knowledge and education', *Journal of Philosophy of Education*, vol. 27, no. 1, pp. 57–68.
Leicester, M 1994, 'Competence, knowledge and education: Reply to Hyland', *Journal of Philosophy of Education*, vol. 28, no. 1, pp. 113–18.
National Training Board 1992, *Policy and guidelines*, (second edition), National Training Board, Canberra.
Oakeshott, M 1962, *Rationalism in politics: And other essays*, Methuen, London.
Schön, D A 1983, *The reflective practitioner: How professionals think in action*, Basic Books, New York.
Spectrum Research and Training 1994, *Designing assessment systems*. A report commissioned by the National Office Skills Formation Advisory Body. Melbourne.
Walker, J C 1992, *A general rationale and conceptual approach to the application of competency-based standards to teaching*, National Project on the Quality of Teaching and Learning, Canberra.
Wolf, A 1989, 'Can competence and knowledge mix?' in Burke, J W (ed.) *competency-based education and training*, The Falmer Press, London.
Wolf, A 1995, *Competence-based assessment*, Open University Press, Buckingham/Philadelphia.

Author

Dr Paul Hager is an associate professor in the School of Adult Education at the University of Technology, Sydney. His research interests include philosophy of adult education and vocational education and training, critical thinking and professional competence. He has published a range of books and research papers on these topics and has been engaged as a consultant by various professions and industry groups in Australasia, notably in the development and implementation of workplace competency-based assessment systems. Currently he is carrying out research on workplace learning and generic competencies.

Assessment at higher levels

Paul Hager and Shelley Gillis

Introduction

'Competency-based assessment at higher levels' refers to the assessment of the more generic, less specific, higher order competencies needed for successful performance in professional and skilled work. Such higher order competencies include things like establishing rapport with a client, thinking critically, acting creatively, solving non-routine problems, making decisions, reflecting on performance, etc. Typically, higher order competencies relate to less routine aspects of work where the worker is called on to frame a skilled and appropriate response to an (often) unique situation.

The competency standards for many occupations specify such 'situational understanding', i.e. they include the idea that workers take account of the effects of the changing contexts in which they are operating. These higher order competencies are not confined to jobs at higher levels of the Australian Standards Framework (ASF), though their importance increases at these levels. Because higher order competencies are significant in the less routine aspects of all jobs, it is important that they figure in competency standards and, hence, in competency-based assessment. It is also evident that higher order competencies, as they have been defined above, overlap with, but are not exhausted by, the Finn-Mayer key competencies. This is consistent with the view that key competencies occur throughout the ASF levels, becoming more complex at the higher levels.

This paper will clarify firstly the important notion of 'levels of competence'. Secondly, the idea of higher order competencies and their relation to levels of competence will be explained with examples. Thirdly, the importance of an holistic and integrated approach to the assessment of higher order competencies will be discussed. Fourthly, the basic principles for assessment of higher order competencies will be outlined and illustrated with examples. Finally, some common issues of concern in the assessment of higher order competencies, such as costs and the assessment of attitudes, will be discussed.

Levels of competence

Competence is an attribute or quality of a person that cannot itself be observed directly. It can only be attributed to individuals on the basis of their performances on real, complex occupational tasks. Determining competence in a occupational area requires a holistic approach to assessment, based on professional judgement (Hager, Athanasou & Gonczi 1994).

This approach to assessment requires competence to be understood as the ability to use and integrate a variety of knowledge and skills to solve real workplace problems (McCurry 1992).

Because competence is something that can only be attributed to individuals on the basis of their performance on real occupational tasks, it follows that employees can have varying levels of competence throughout their working career. For instance, the level of competence for an entry-level employee would be lower than that of an expert in the same occupation. Hence, an employee's competency level can be thought of as falling at some place on a continuum. As the employee's proficiency increases, we infer that their competence has moved to a higher level on the continuum. Thus, assessment of competence should emphasise and encourage the occupational growth of employees and quality improvement in the workplace. The development of competence is an important component of the training reform agenda in that it promotes lifelong learning and motivates employees to achieve their full potential.

By conceptualising competence as a continuum, lower and higher levels of competence can be established for the same occupational area across different ASF levels. It is this notion of competence as 'scales of proficiency' that enabled the development of the ASF levels:

> *The levels in the ASF are points in a continuum of competencies, and as such simply reflect the needs of the Australian environment.*
>
> (National Training Board, 1992)

Thus, the various levels of competence within an occupational stream are reflected in their aligned ASF levels. The National Warehousing and Distribution Skills Pty Ltd competency standards can be used to illustrate the growth of competence. The following examples taken from the warehousing and distribution competency standards reflect the increasing proficiency required to provide customer service as an employee progresses from level 1 through to levels 3 and 5. The performance criteria and evidence guide spell out what is involved in providing competent customer service at each of these three ASF levels.

Competency stream—customer service

ASF level	Definition	Element	Performance criteria	Evidence guide
Level 5	Work is performed under broad guidance and involves responsibility for planning and managing the work of others. Competencies are applied to both routine and non-routine procedures. Substantial judgement is involved. A variety of skills and depth of knowledge are applied to a wide range of tasks.	5.3.1 Plan and manage the operation of the warehouse taking customer needs into account.	• Thorough knowledge of customer needs is demonstrated. • Regular discussion is undertaken with customers. • Plans are communicated to management and staff. • Operations are adjusted to meet customer needs.	The ability to plan and manage the activities of the warehouse to meet customer needs.
Level 3	Work at this level is performed under limited supervision. Team organisation may be involved. The range of tasks performed is broad with some complexity in the extent and choice of action required.	3.3.1 Practice and promote customer service.	• A helpful attitude is displayed when responding to enquiries. • Information being given to the customer is up-to-date. • Customer is questioned in a way that displays an understanding of needs. • Response to the enquiry is put into action as completely and accurately as possible. • Checks are made to ensure that the customer's needs have been met.	The ability to respond to and resolve customer enquiries.
Level 1	Work at this level is performed under constant supervision. Knowledge and skills are applied to a limited range of tasks and roles. Limited choice and judgement is required within established routines and procedures.	1.3.1 Receive internal customer enquiries.	• Customer requirements are identified. • Information is sought from customers about their needs. • Enquiries are reported. • Resolutions are received and communicated.	The ability to receive internal customer enquiries quickly and effectively.

Taken from the *National Warehousing and Distribution Competency Standards* (National Warehousing and Distribution Skills Pty Ltd 1993).

Higher order competencies

As was noted at the start of this paper, higher order competencies typically relate to less routine aspects of work. Though these higher order competencies are not confined to jobs at higher levels of the ASF Framework, their importance increases at these levels. This can be seen from the above examples taken from the warehousing and distribution competency standards. At level 1 of customer service, the emphasis is on the worker following routine procedures and exercising a limited amount of judgement.

At level 5 of customer service, the emphasis is on the worker applying non-routine as well as routine procedures and exercising a substantial amount of judgement. Hence, an employee who progresses from ASF level 1 to level 5 in this occupation, needs, among other things, to go through a developmental process in customer service competence in which greater proficiency is acquired in such higher order competencies as decision making, reflecting on performance, thinking critically and problem solving.

The focus of this paper is on how to ensure that assessment of these higher order competencies is successfully built into competency-based assessments. The ways that competence is assessed are strongly influenced by how we think about the notion of competence. The kind of competency-based assessment that most successfully measures higher order competencies is based on an holistic and integrated view of competence. However this view is not evident in much of the literature on the subject. In fact, there are three distinct views of the nature of competence. The view of competence that is adopted will make a big difference to the assessment process that is selected.

Views of competence and how best to assess it

Behavioural approach

The first, and probably the most widely held, view of the nature of competence is task-based or behaviourist. In it, competence is thought of in terms of the separate activities associated with the completion of independent tasks. The assembly line represents a classic case of this approach to describing work. Work is broken down into a series of relatively simple, repetitive tasks with each worker performing their assigned task in a standard unchanging way as each new item arrives at their station on the assembly line. Here, the aim is to specify the competencies clearly so that there can be no disagreement about what constitutes satisfactory performance.

The task becomes the competency, so that, for example, a worker who can mount a fuel pump on an engine chassis correctly is said to possess the competency of fuel pump mounting. This approach to describing work ignores the connections between tasks and thereby the possibility that some competencies relate to whole sets of tasks. Evidence for the possession of the competency in this model is based on direct observation of performance. Hence, assessment is seen in terms of direct observation and marking off items on a checklist.

Clearly, such an approach is not well suited to the description of the various higher order competencies, which underpin broader rather than very narrow aspects of performance. For example, a worker at level 5 of the customer service stream of the

warehousing and distribution competency standards is required to 'Plan and manage the operation of the warehouse, taking customer needs into account'. Not only does competent performance of this task call for a complex combination of higher order competencies, the assessment of performance on this element would obviously call for an holistic judgement based on evidence gathered from a variety of appropriate sources. Direct observation and ticking off items on a checklist would not be the appropriate method of assessing of competence at this higher level and would certainly compromise the validity of the assessment. In fact, such an approach to competency-based assessment, particularly at the higher ASF levels, would reduce the assessment to such an extent it would have minimal, if any benefits to the stakeholders (e.g. employees, employers, trainers).

However, because this behaviourist approach promises simplicity and clarity, it was adopted by many of the industries in Australia and England when competency standards were first developed. It is usually the model in people's minds when they attack competency-based assessment (Ashworth and Saxton 1990; Field 1991). As these authors and others have pointed out, the weaknesses of this approach are easy to identify.

Among other things, it neglects underlying attributes that are crucial to performance; it ignores group processes and their effect on performance; it overlooks the complexity of performance in the real world; and it misses the role of professional judgement in intelligent performance (Preston & Walker 1993). At best, this approach is suited to describing repetitive work of a routine kind where there are no significant higher order competencies involved. Clearly it is inappropriate for understanding and assessing complex work activities and, given the trend to more complex workplaces, there are very serious doubts about its relevance to current work at any ASF level.

Generic approach

The second approach to the nature of competence concentrates on the most general abilities of the worker that are crucial to effective performance. For example, turning again to level 5 of the customer service stream of the warehousing and distribution competency standards, it is clear that competent performance requires, amongst other things, knowledge (of customer needs), communication, planning and managing.

According to this view of competence a generic ability, such as communication can be applied to many or all relevant situations. Thus, competencies are thought of as general attributes, ignoring the context in which they might be applied. This approach has been popular particularly in management and related occupations (McBer 1978; Boyatzis 1982). It can be described as a reaction to the omission of higher order competencies in the behaviourist approach, by describing work as nothing but higher order competencies. Thus, it is an attempt to avoid the shortcomings of the behaviourist view.

On this view of competence, assessment becomes the assessment of a series of generic abilities such as communication, planning and management. The assumption is that anyone who has these capabilities will be able to exercise them in any situation that calls for them.

There is, however, a well-known problem with this approach. There is no certainty that generic competencies actually exist independently of a context. The evidence from the novice/expert research literature (Chi, Feltovitch & Glaser 1981; Greeno 1989; McGaw

1993) and from related literature (Townsend 1993) suggests that expertise (which can be characterised as high levels of competence in an occupational area) is significantly domain specific. That is, individuals demonstrate little capacity to transfer expertise from one occupation to another, unless the two occupations are closely related.

It is even possible that contexts within the one industry or occupation could vary so much that competence in one context would not guarantee competence in others. For example, let us return yet again to the worker at level 5 of the customer service stream of the warehousing and distribution competency standards, who is required to 'Plan and manage the operation of the warehouse taking customer needs into account'. It could be the case that someone is extremely competent at this work function when the customers generally are third-generation Australians, but not competent when the customers are much more ethnically diverse.

Traditionally, vocational education and training courses have concentrated on developing and assessing generic skills in the broad context of an occupation, inferring that these are the basis of successful practice in the future. It would seem that effective training in and assessment of the higher levels of competence should occur in situations that closely resemble the kinds of contexts in which the higher order competencies occur in the practice of the particular occupation.

Integrated approach

The third view of competence seeks to integrate general attributes with the context in which these attributes will be employed. This approach looks at the complex combinations of attributes (knowledge, attitudes, values and skills) which are used to understand and function within the particular situation in which workers find themselves. Competence is seen as relational. That is, it brings together two different things: the abilities of individuals (deriving from combinations of attributes) and the tasks that need to be performed in particular situations. Thus, competence is viewed as a complex structuring of attributes needed for intelligent performance in specific situations.

This approach has been called the 'integrated' or holistic approach to competence (Gonczi, Hager & Oliver 1990). It is the view that has been adopted by many industries and occupations in Australia. It is also the conception of competence that best accords with the official view of competence:

> *Competency standards provide the specification of the knowledge and skill and the application of that knowledge and skill to the standard of performance required in employment.*
>
> (National Training Board 1992)

The National Training Board policy documents are clear that the concept of competence is intended to capture broader aspects of work performance including:

- performing at an acceptable level of technical skill;
- organising one's tasks;
- responding and reacting appropriately when things go wrong;

- fulfilling a role in the scheme of things at work; and
- transferring skills and knowledge to new situations.

The integrated approach acknowledges competence as a developmental process or continuum that promotes occupational growth. Furthermore, it is plausible to suggest that this approach to competency-based assessment overcomes all of the significant objections to the competency standards that have been identified in the literature (Hager 1994). It allows for, among other things, such higher level considerations as the role of ethics and values in competent performance, the need for reflective practice, the importance of diverse contexts and the fact that, in many cases, there is more than one way of performing competently. Competency-based assessment based on the integrated approach is the topic of the next section.

Some basic principles for assessment at higher levels of competence

Integrated or holistic competency-based assessment should be used wherever possible to cover multiple elements and/or units of the competency standards. This gives the assessment validity, whilst at the same time it saves costs and maximises efficiency. This equally applies to assessments at lower levels of competence, but it is essential for assessment of higher order competencies since these typically are present only in the performance of relatively complex and demanding work activities.

Not only is integrated holistic assessment of higher order competencies more valid than alternatives, experience (e.g. the law example discussed below) suggests that it also appears to be very sound to those being assessed. This should serve to minimise possible implementation difficulties. As more competency-based assessment is used, the identification of examples of best practice in integrated holistic assessment will be an important topic for further research.

When developing an holistic assessment, the first step is to determine which assessment methods are most capable of assessing competence in an integrated manner. Integrated approaches seek to combine knowledge, understanding, problem solving, technical skills, attitudes and ethics in assessment. Integrated or holistic assessment can be characterised as:

- problem oriented;
- interdisciplinary;
- embracing actual practice;
- covering groups of competencies;
- focussing on common circumstances;
- demanding analytical abilities; and
- combining theory and practice.

The integration is made possible by methods which assess a number of elements and all their performance criteria simultaneously. Thus carefully planned simulations enable various elements of the competency standards to be assessed (or partly assessed) in the one assessment event.

Likewise, standardised cases can be used as integrated assessment events to assess several elements simultaneously, such as conducting individual interviews, monitoring the progress of clients, and compiling case records and reports. These standardised cases are also capable of providing evidence of a number of attributes at the same time, such as attitudes, communication skills and background knowledge.

Example

A specific example of holistic assessment with a focus on higher order competencies is the following from family law. Since 1993, the Law Society of NSW has created a classification, 'specialist' lawyer, in a series of specialisations. They have developed integrated competency standards for each specialisation, and have devised and implemented in each case an holistic assessment strategy based on the competency standards.

One of the assessment events requires candidates to conduct a simulated interview with a person acting in the role of a client. This simulation, of approximately 60 minutes duration, is videotaped and the videotape assessed by the examiners. The various versions of this simulation have been constructed so as to assess a wide range of the competency standards including those related to interaction between the solicitor and the client, taking instructions and giving advice, assessing facts and legal options, canvassing the options with the client, and developing the initial plan.

Higher order competencies tested by the simulation include interpersonal communication, critical thinking, decision making, planning, evidence gathering and acting ethically, as well as various kinds of knowledge and their application. Thus, the simulation was developed with elements and performance criteria from a variety of units in mind. While depicting different types of clients and different situations and problems, the six simulation exercises developed all have the same basic structure:

- an immediate need;
- long term issues;
- information not disclosed unless appropriate questions are asked;
- client's hidden agenda;
- presentation of a problem that requires non-legal solutions (including some of a religious/cultural kind that require sensitive handling);
- an ethical issue; and
- a query on costs.

It is true, in general, that assessment of performance of real work situations is more valid than that of simulated work situations. However, practical and ethical considerations, the relative weight of which varies with the nature of the occupation, can sometimes tip the scales in favour of simulations. In the medical and paramedical occupations, for example, simulated patients are likely to be much less satisfactory than real patients for purposes of performance assessment. However, in the legal area, simulated clients have distinct advantages. Provided that the person playing the role of the client has been well trained, it is possible to use carefully designed cases that are much richer from an assessment point of view than are typical real cases. Thus a well-designed simulation can readily yield information which would be more difficult to obtain from an appropriate combination of real cases.

The importance of context

The crucial role of context in assessing higher order competencies has already been emphasised. While generic descriptions such as critical thinking, problem solving and communicating have their conceptual uses, these competencies preferably should be assessed in a situation that resembles a real work context. Not only is this because problem solving, for example, might take very different forms in different occupations, there also may well be significant differences within a single occupation. As Sheldrake (1995) puts it:

> . . . *generic management competencies comprise the general set of skills, knowledge and application associated with the management function; however, they may well have to be specifically varied in a particular place. (p. 3)*

The earlier example, involving the warehousing and distribution competency standards where in one case the customers generally were third-generation Australians, but in another were much more ethnically diverse, illustrates the importance of valid assessment needing to take account of differing contexts.

Choice of assessment method

As a general rule the assessment method should be matched with the type of performance being assessed. For example, in medicine, patient diagnosis probably is best assessed by direct observation of performance using a real or simulated patient, whereas more complex competencies, such as interpretation of a pathology report, can best be assessed more indirectly through some kind of questioning situation such as a written or oral test. Thus, we should always try to select the methods that are most relevant to the kind of performance being assessed.

Having selected a method, we should look to build into the assessment a wide variety of the higher order competencies that need to be assessed. The family law simulation discussed above illustrated this point. In the case of family law, three methods proved to be sufficient for assessing candidates' competence. The three methods were the simulation, a knowledge exam, and the completion of a mock file.

The methods available for competency-based assessment include:

Questioning techniques—projects/assignments; problems; case studies; written tests

Simulations—simulated clients/customers; simulated workplaces; situational exercises

Skills tests—standardised clients/customers; work samples; role plays

Direct observation—supervisor evaluations; practice year; professional year; internship; industrial experience

Evidence of prior learning—portfolios; log books; qualifications; referees; records of achievement.

Breadth of evidence base

A general principle, important for ensuring the validity of competency-based assessment, is that the narrower the base of evidence for determining competence, the less possible it will be to generalise about the performance of other tasks. An obvious example is that performance on paper-and-pencil tests alone would probably be too narrow a base for assessing competence in any occupation. Equally, given the significance of higher order competencies in many occupations, observation of performance on a small range of routine tasks would be too narrow a base for assessing competence. Hence, it is recommended that a combination of the methods listed above be used for providing evidence on which to determine competence. In choosing this mixture of methods an important consideration will be how well the various relevant higher order competencies are being assessed.

The inference of competence

One important consequence of adopting the integrated approach is that competence is not something that is observed directly. Rather, competence is inferred from performance. This means that, as for other kinds of assessment, there are procedures available which will maximise the validity and reliability of assessments of competence. If these procedures are followed then assessment of competence is as objective as any of the alternatives.

Those who criticise competency-based assessment on the grounds that it involves inference apparently are unaware that this is the case for all typical assessment methods. Tests of knowledge, for example, usually sample a small fraction of the required knowledge, then on the basis of a score that is typically well below 100 per cent, an inference is made as to whether or not the student knows enough to be satisfactory (Wolf 1989).

It is in the nature of higher order competencies, such as communicating, planning, managing, etc. that assessment will sample at best only some of the instances in which these competencies are required in the workplace. Hence, competency-based assessment, in common with other types of assessment, involves inference that is subject to error. In all cases, the way to proceed is to gather the kind of evidence that will make the inference valid. This means following available procedures to ensure that inferences about competence are sound (Gonczi, Hager and Athanasou 1993).

Problems with assessing attitudes/values

Although attitudes and values are important in all occupations, they are particularly significant in occupations at the higher ASF levels. For instance, at the higher ASF levels it is more likely that individuals are responsible for the well being of other workers, and there may be a requirement for adherence to codes of conduct. Research in the behavioural sciences has discovered the difficulties of conducting reliable assessments of attitudes and values in isolation (Groth-Marnat 1990; Oppenheim 1966). The most common problem with assessing attitudes is the tendency for individuals to provide responses that reflect socially desirable attitudes and not necessarily their true beliefs and feelings. The motivation to supply socially desirable responses is particularly significant when there are incentives associated with the outcomes of the

assessment (e.g. promotions, pay structure). The difficulties associated with assessing attitudes have led some people to argue that competency assessments should be restricted to skills and knowledge (i.e., attributes that are easily observable). Excluding attitudes and values that are specified within the performance criteria would, however, undermine validity.

To ensure that assessments are both valid and reliable, an integrated approach to assessment, using multiple methods of assessment (as described above), needs to be adopted. This holistic approach to assessment would enable attitudes that elicit desired behaviour patterns to be inferred from measurements of actual or simulated work performance. Such an approach to assessing attitudes also has the advantage of allaying industrial fears about possible misuse of attitude assessment (see the companion paper *Assessment of knowledge and attitudes in competency-based vocational education and training).*

Costs

There is no empirical research that has examined costs specific to assessment at higher levels of competence. Irrespective of what level of competence one is assessing, there will be direct costs associated with the development of assessment materials, administration (e.g. obtaining assessment instruments, hiring facilities/equipment) and recording and reporting the assessment (Gillis, Chaplin & Griffin 1995; McDonald 1995). Although there do not appear to be any costs that are specific to higher levels of competence, it could be argued that assessments may be more expensive at the higher ASF levels for the following two reasons:

- money loss in productive work time of both the assessor and assessee (hourly pay rate would be higher at the higher ASF levels); and
- limited availability of assessors with the appropriate vocational competence to assess (higher levels of competence may have to be assessed by external assessors).

However, the actual direct costs will depend on the type of assessment being conducted and the desired quality of the assessment, irrespective of the ASF levels of the assessees. Costs obviously are going to be great if the assessment comprises extensive and rigorous testing that minimises reliance on professional judgement and inference. As such, it is argued that holistic assessments that measure a number of elements and all their performance criteria simultaneously are the most cost-effective approach to competency-based assessment at all levels of competence.

Conclusion

Assessment at higher levels of competence requires an integrated approach to assessment that is context dependent, recognises competence as a developmental process and uses professional judgement. Although an integrated approach to assessment is recommended for assessment at all levels of competence, this approach is essential to the assessment of higher order competencies such as problem solving, communication skills, management skills and attitudes. Future research needs to be conducted to identify examples of best practice in integrated holistic assessments as well as to determine costs that are specific to assessment at higher levels of competence.

Bibliography

Ashworth, P D & Saxton, J 1990, 'On "competence,"' *Journal of Further and Higher Education*, vol. 14, pp. 3–25.

Boyatzis, R 1982, *The competent manager*, Wiley, New York.

Chi, M T H, Feltovitch, P & Glaser, R 1981, 'Categorisation and representation of physics problems by experts and novices,' *Cognitive Sciences*, vol. 5, pp. 121–151.

Field, J 1991, 'Competency and the pedagogy of labour,' *Studies in the Education of Adults*, vol. 23, no. 1, pp. 41–52.

Gillis, S, Chaplin, S, & Griffin, P 1995, *Industry Skills Assessment and Recognition*, Royal Melbourne Institute of Technology, Melbourne.

Gonczi, A, Hager, P & Oliver, L 1990, 'Establishing competency-based standards in the professions', research paper no. 1, National Office of Overseas Skills Recognition, Department of Employment, Education and Training, AGPS, Canberra.

Gonczi, A, Hager, P & Athanasou, J 1993, 'The development of competency-based assessment strategies for the professions,' research paper no. 8, National Office of Overseas Skills Recognition, Department of Employment, Education and Training, AGPS, Canberra.

Greeno, J 1989, 'A perspective on thinking,' *American Psychologist*, vol. 44, no. 2.

Groth-Marnat, G 1990, *Handbook of psychological assessment*, second edition, Wiley, New York.

Hager, P 1994, 'Is there a cogent philosophical argument against competency standards?', *Australian Journal of Education*, vol. 38, no. 1, pp. 3–18.

Hager, P, Athanasou, J & Gonczi, A 1994, *Assessment technical manual*. AGPS, Canberra.

Hager, P & Gonczi, A 1993, 'Attributes and competence,' *Australian and New Zealand Journal of Vocational Education Research*, vol. 1, no. 1, pp. 36–45.

Hager, P, Gonczi, A & Athanasou, J 1994, 'General issues about assessment of competence,' *Assessment & Evaluation in Higher Education*, vol. 19, no. 1, pp. 3–16.

Hall, W C 1994, *Competency-based training and assessment*, National Centre for Vocational Education Research, Adelaide.

McBer & Company 1978, *Understanding of competence*, McBer, Boston.

McCurry, D 1994, 'Assessing standards of competence,' In Gonczi, A (ed.), *Developing a competent workforce: Adult learning strategies for vocational educators and trainers*, National Centre for Vocational Education Research Ltd, Adelaide.

McDonald, R 1995, *Costing training and assessment—issues paper*. Assessment Centre for Vocational Education, NSW TAFE, Sydney.

McGaw, B 1993, 'Competency-based assessment: Measurement issues', National Assessment Research Forum. NSW TAFE Commission, Sydney.

National Training Board 1992, *Policy and guidelines*, second edition, National Training Board, Canberra.

National Warehousing and Distribution Skills Pty Ltd 1993, *National warehousing and distribution competency standards*, National Warehousing and Distribution Skills Pty Ltd, Melbourne.

Oppenheim, A N 1966, *Questionnaire design and attitude measurement*, Basic Books Inc., New York.

Preston, B & Walker, J C 1993, 'Competency standards in the professions and higher education: A holistic approach,' in Collins, C (ed.), *Competencies: The competencies debate in Australian education and training*, Australian College of Education, Canberra.

Sheldrake, P 1995, 'Where do competencies fit in management?—The implications of the development and implementation of management competency standards,' paper presented to the Global Competencies—Workplace Outcomes Conference, Darling Harbour Convention Centre, Sydney, 3 March 1995.

Townsend, M A R 1993, 'Generic skills and transfer,' unpublished report commissioned by the Policy, Research and Development Division of the New Zealand Qualifications Authority.

Wolf, A 1989, 'Can competence and knowledge mix?,' in Burke, J W (ed.), *Competency-based education and training*, The Falmer Press, London.

Wolf, A 1995, *Competence-based assessment*, Open University Press, Buckingham/ Philadelphia.

Authors

Dr Paul Hager is an associate professor in the School of Adult Education at the University of Technology, Sydney. His research interests include philosophy of adult education and vocational education and training, critical thinking and professional competence. He has published a range of books and research papers on these topics and has been engaged as a consultant by various professions and industry groups in Australasia, notably in the development and implementation of workplace competency-based assessment systems. Currently he is carrying out research on workplace learning and generic competencies.

Shelley Gillis, BBSc, Grad Dip Applied Psych, is a researcher at the Assessment Research Centre, RMIT, Melbourne. Her major research activities are in the area of competency-based assessment and system design. Currently undertaking her PhD studies in the field of competency-based assessment, her research examines competency acquisition and maintenance of higher order competencies as a function of training and experience at the higher levels of the Australian Standards Framework. Her professional experience with industry training advisory bodies includes the design of standards, the development of assessment protocols and the establishment and management of recording and monitoring systems. She is also a certified workplace assessor trainer for the Assessment Research Centre which is a registered private provider of assessor training. Prior to joining the Assessment Research Centre, Shelley Gillis worked with the National Road Transport Industry Training Council as a research assistant.

Self-assessment and peer-assessment

Geof Hawke and Patrick Griffin

Introduction

The use of self-assessment and peer-assessment in competency-based vocational education and training has been much discussed but there have been few attempts to implement it in any formal way, either in Australia or overseas.

While the two different approaches that are discussed in this paper can operate separately from each other, in practice they are often used together. In particular, there are almost no examples of peer-assessment being used without some form of self-assessment also being applied.

This does not mean that there is little use or knowledge of these two forms of assessment. Indeed, there is substantial literature on both. However, in general, both have been used as approaches to teaching and learning rather than as formal assessment methods in themselves. Readers who are interested in learning more about self- and peer-assessment for these purposes can refer to Boud (1980) or, when available, Boud (in press).

What is self-assessment?

Self-assessment in a competency-based system is a process in which an individual gathers evidence of their own competence and makes a decision about whether competence has been achieved by referring to the competency standards defined for a particular industry.

It is a continuous, holistic process in which the individual draws on the whole of their experience and understanding in gathering and interpreting the assessment information. While it is clear that individuals can, given adequate benchmarks or criteria, reflect upon and monitor their own improvement, questions remain as to the capacity of self-assessments to provide valid and reliable outcomes. We will discuss this in more detail at a later stage.

Experience in using self-assessment, especially in a competency-based context, is still uncommon and much of this review will rely on its use in other situations.

What is peer-assessment?

Peer-assessment, like self-assessment, is not fundamentally different from other forms of assessment. Its unique feature is the relationship of the assessor to the assessee. This relationship is the significant aspect of peer-assessment as it creates both the strengths and weaknesses of the approach.

Peer-assessment is the process of using information and judgements from colleagues or peers in determining some aspect of the competence of an individual. It is used to provide feedback on an individual's learning and to make judgements about whether or not competence has been achieved.

In some cases being competent in a particular workplace role entails developing competence to assess colleagues. In such cases peer-assessment can be an educational outcome as well as a means of judging achievement of competence.

Who uses self- and peer-assessment information?

How (or, indeed, whether) self- or peer-assessment is used depends on the purpose for which the assessment is undertaken. Therefore, it is important that the audience for the information is identified.

Where the primary audience is the individual learner, self- and peer-assessment serves essentially a 'formative' purpose in supplementing and extending the range of feedback the learner receives from other sources.

'Formative' assessment is the use of assessment in a learning process to inform learners about either the extent to which they are progressing towards competence and/or those areas which require further development. In competency-based education and training, this occurs throughout the learning process.

Toop (1995) argues strongly for self-assessment to be used to provide a screening process by which learners are able to assess their readiness for a more formal assessment to be conducted by others.

Where the primary audience is a certifying, licensing or similar authority, or an employer or provider who wishes to know if competence has been achieved, then self-assessment is used in a summative way, i.e. it makes a final judgement as to whether or not the learner has achieved the outcomes defined for that unit of learning. It is vital, then, that the system of self- or peer-assessment meets the standard criteria of validity, reliability, flexibility and fairness (Hager, Athanasou and Gonczi 1994).

Employment-related uses

Self-assessment has not been employed widely in industry or for employment-related purposes. Clearly, the use of self-assessment for promotion or selection purposes is unlikely to be acceptable at any stage.

However, Bennet (1992) has argued that self-assessment can contribute to the validity and reliability of workplace assessment. He believes that, once assessment criteria within endorsed competency standards are clearly defined, all assessment methods, including self-assessment, will be more efficient. In the case of self-assessment, he believes that this will occur if the learner has internalised the standards and criteria. This requires the learner to 'interact genuinely with their workplace supervisors and disclose to them their thoughts, feelings and intentions about their work'. It seems unlikely that this could easily be achieved in most contemporary work environments.

The literature shows that peer-assessment has been used primarily as an adjunct to self-assessment: it is seen as providing the assessee with a further source of information and a different frame of reference in which to judge evidence of competence. There are, however, examples of peer-assessment being used as a technique in its own right though these generally have been restricted to professional quality assurance systems. Examples can be found in accountancy, engineering, medicine, community work, higher education and science.

Professional registration-quality assurance is probably the area which uses peer-assessment most often. Frequently called 'peer review', many professional associations—especially in the USA—use peer-assessments for ongoing assessment of the eligibility of professionals to retain their membership or registration. In other related cases, it is part of a profession's internal quality assurance processes (including so-called self-regulation). Examples of this can be found in accounting (Patillo 1984; McCabe 1993), medicine (Peer Review Resources Centre 1993), engineering (American Consulting Engineers Council—American Society of Civil Engineers 1990) and higher education (Boud 1980; Goedegbuure 1990).

In all these cases, peer judgements play a role in determining the ongoing professional status of individuals or groups. Generally, this information is used to shape change and to modify behaviours. Only rarely is peer judgement directly used to decide the careers of individuals.

In particular, there are no reported Australian examples of peer-assessment playing a direct role in employment or promotion. However there is no reason, beyond perceptions of its acceptability, why it should not be used for such purposes. However, an assessment process which is not acceptable to key players is not a workable system. Therefore, introducing peer-assessment requires the support and agreement of all the interested parties, and this entails guaranteeing a level of certainty in the fairness of the process.

Relatively little work has been undertaken in the use of peer-assessment in the recognition of competence. The notable exception is the use of references and related documents as part of the recognition of prior learning. Some Australian examples of this use include Dugas & Horrocks (1993), Holland (1993).

The work of Kearney (1993) and of the National Office Skills Formation Advisory Body (Fitzherbert 1993) are the only other significant sources of experience. In the latter case, a substantial, structured system of assessing competence for administrative and clerical workers has been developed that relies heavily on both self- and peer-assessment.

Uses in education/training

Most of the literature on self-assessment can be found in schools, colleges and higher education settings.

For instance, Broadfoot (1989) argues that the use of self-assessment in the UK has increased the quality of learning. However, the paper provides little evidence to support this view beyond the students' own self evaluations.

Boud (1980; in press) however, provides a comprehensive analysis of the uses of self-assessment within higher education. In the main, the learners use their own evaluation of their progress to inform and modify their future learning.

The role of self- and peer-assessment within education or training programs often goes beyond the purpose of evaluation alone to include a specifically educational purpose.

For example, competency standards in many industry areas expect that workers at various levels will be competent in assessing their own performance and/or the performance of others. In the case of self-assessment, the growing emphasis on quality assurance and the responsibility of each employee for the quality of their own work has made this a critical area of skill development. Peer-assessment occurs as part of a supervisory or management role, while at other times it is a natural result of the team-oriented nature of the work itself, for example, in community health (Northcote District Health Council 1988). In these cases, education and training programs may incorporate the development of self- or peer-assessment skills as part of the teaching/learning process itself (Kearney 1993).

The work of Kearney (1992) is also relevant. Kearney has developed an approach he describes as 'collaborative assessment' in which trainees actively participate in their own assessment and in that of other trainees in the group. Kearney provides one of the most comprehensive analyses of the use of both self- and peer-assessment. It outlines numerous approaches to collaborative assessment and provides a range of exercises which learners should practice before using self- or peer-assessment. For Kearney, the great value of this approach is that it 'puts a human face' on what is otherwise a somewhat difficult undertaking.

Peer-assessment as part of the classroom learning process is used in all sectors of the educational system—schools, TAFE and higher education. It is also a critical component of much work-based training.

Its use in off-the-job education and training varies from using peers to mark assessments or tests to building peer feedback into group projects, classroom discussions and the like. In the on-the-job context, peer-assessment often forms a key part of the total learning environment. This occurs most often in enterprises that are adopting team-based approaches to work organisation. In these organisations, the team frequently is expected to operate within a climate of continuous improvement in which positive, constructive feedback on performance of both individuals and the group is an essential component. In this case, the assessment usually is given orally and is not recorded.

Why use self- or peer-assessment?

Self- and peer-assessment provide a way of gathering evidence in a competency-based system. Unlike most other forms of evidence-gathering such as direct observation, formal tests/measurements, self- and peer-assessment are based on the judgements of other people who are not formal assessors.

It is clear that the understanding of the learner of his or her own strengths and weaknesses plays a critical role in their capacity to use and benefit from an ongoing learning situation. As a summative assessment however, there are good grounds to regard self-assessment as an approach only to be used with caution.

Most commentators (National Training Board 1991; Cohen 1993; Raffe 1992) have emphasised the importance of using expert assessors to manage formal, summative assessment. Because assessees are also learners, this cannot occur in self-assessment.

The capacity of the learner to comprehend fully the standards and assessment criteria is a significant issue in determining whether or not this is a valid and effective method.

Also, while the literature reports a range of findings, the general view is that self-assessments show systematic bias. Learners typically underestimate their performance, and even in circumstances where this is not so (Griffin 1988) the level of consistency between self-assessment and expert judgements is generally poor.

Peer-assessment is also affected by subjective factors. Judgements are coloured by the relationships that exist between the assessee and their peers. This is both its strength and its weakness.

In work situations, the interaction of an individual with their peers influences significantly the effectiveness of each as a worker. Assessment which draws on the perspectives of peers then provides a highly realistic and usually holistic assessment of performance or competence. Moreover, it is one which draws on the peers' interactions over time rather than in just a few contrived assessment situations.

At the same time, there is considerable evidence that such judgements are extremely variable in their quality. Politeness, embarrassment, bias, prejudice and simple misunderstandings are all highly probable in such an arrangement. Research indicates that the quality of information obtained from peer-assessment heavily depends on the structure and clarity of the instructions given to the assessors. For example, an instruction to 'evaluate the person's ability to interact with others' provides little clear guidance as to the types of interaction, the context(s) within which interaction is to be judged, the range of people with whom interaction is to occur, and also does not indicate the features of the interaction or the standards of performance against which the person's competence is to be judged.

Peer-assessment provides another important view of the worker's competence often not available from other sources. The peer typically sees the worker most or every day, performing in a wider range of situations than would be possible in more constructed assessment situations. Thus, the peer has access to information on the worker's performance over time, their consistency and their capacity to perform in unusual situations. For those skill areas that draw on complex aspects of workplace competence, the more detailed knowledge of performance that a peer can provide is valuable and, sometimes, essential.

The assessment process in self- and peer-assessment

Formative

Formative assessment occurs during training, whereas summative assessment occurs at the end. The clerical and administrative assessment project (Fitzherbert 1993; National Office Skills Formation Advisory Body 1993) provides a good example of the use of self- and peer-assessment in a competency-based context.

The National Office Skills Formation Advisory Body (NOSFAB) is responsible for the development of competency standards for the non-government clerical and administrative workforce. To assist the implementation and use of those standards and to promote and expand training opportunities, NOSFAB has developed a comprehensive system of workplace training and assessment. Self- and peer-assessment is a key tool in the overall system of assessment they have adopted.

While the final judgement of competence is the responsibility of a certified assessor (often the supervisor), self- and peer-assessment has been used to provide the learner with feedback on their progress and to identify when learners are properly prepared for their formal assessment. The system uses a fairly detailed and clearly defined set of documentation to structure the judgements provided by learners, their peers and their supervisors. The endorsed competency standards are used as the reference point and a record book. The book records completion of each of the units of competency and provides guidance on assessment. The elements, range and performance criteria in the standards provide the learner, peer and workplace assessor with detailed information on the context of the assessments and the criteria which characterise successful performance.

The system has proved to be both simple and effective and is used widely. Anecdotal evidence suggests that supporting peer- and self-assessment in this way, moreover, has kept costs within levels employers are prepared to accept.

Summative

Summative assessment occurs at the end of training. As a form or part of summative assessment, self- and peer-assessment is rarely used as the sole source of information.

Indeed, no examples of self-assessment playing a major part in a summative assessment could be found.

In most cases, final judgements about the achievement of competence incorporate peer-assessment as a supplement to information already collected from a range of other sources. An example is the use of peer-assessment as a component of quality assurance systems in higher education (Somervell 1993). In this example, information from peers is one of a series of intersecting views that combine to form an overall judgement.

One exception has sought to use peer-assessment to determine an assessment outcome which reflects an individual's contribution to a group process (such as in a group project). This is a difficult area of assessment practices, but initial results suggest that it is worth pursuing (Conway et al. 1993; Keaten and Richardson 1993).

Many recognition of prior learning (RPL) processes have used peer information as part of a summative evaluation of the applicant. In the main, these have relied on the collection of referees' reports that often have been prepared in the past and for quite different purposes. This creates problems of consistency, and for this reason, some RPL systems now require that portfolios contain current information collected according to a fixed set of criteria for structure and content.

When used in a formal context, peer-assessment almost invariably involves the preparation, by colleagues or associates of the assessee, of a written report. The nature of the report will vary according to the context.

How would you set up a peer-assessment system?

Incorporating peer-assessment requires that the same safeguards and requirements be met as for any other assessment method. The decision criteria set out by Toop, Gibb and Worsnop (1994) form a useful guide in determining whether, and if so where, assessment by peers has a role in an assessment system.

Purpose of assessment

The most fundamental decision to be made involves establishing why you want to include self- or peer-assessment in your overall program.

As has been made clear above, self- and peer-assessment generally has not been a significant feature of summative assessment systems. Rather, it is seen as an important tool within the teaching/learning process. In this context, self- and peer-assessment provides feedback to learners and teachers/trainers and, in increasingly many cases, operates as a learning outcome in its own right.

If this is the principal purpose for using self- and peer-assessment then many of the considerations that follow are of lesser significance as they focus on the more complex issue of self- and peer-assessment as a formal, summative method.

Suitability of the learner or peers as assessors

As has been discussed earlier, there is a general view that assessment in a summative context needs to be managed by assessors who are expert in the area to be assessed and who have skills in the assessment process as well. This is clearly not the case where learners are undertaking self-assessment. If self-assessment is to be used, the learner will require some assessment skills. Moreover, the learner's self-assessment will need to be supported or verified by other reliable sources. In most cases these requirements create difficulties which are complex to resolve and which make this a less attractive option.

Peers usually have the greatest familiarity with the standard of ongoing work performed by their colleagues. In addition, they usually have more current knowledge of actual work practices, and, often have considerable self-interest in ensuring that judgements of the competence of their peers are fair and accurate.

On the other hand, peers are often conservative and unwilling to identify changes in practice that are necessary or desirable, are unwilling to 'dob-in' their mates, and do

not themselves have competencies they are called on to assess.
The capacity of peers to provide sound, valid and reliable judgements is a critical issue. The literature provides no clear guidance. Goldman (1994) in reviewing a number of studies found that, overall, the reliability of peer-assessments was not high. Other studies (Hughes and Large 1993) report high levels of agreement between peer judgements and those from other sources. The difference in outcomes most likely is explained by the high variability in the way peer judgements have been collected and used. However, use of selection criteria such as competence in the work to be assessed, capacity to provide honest, analytical feedback, absence of any relationship which might cloud or bias judgement and good communication skills are likely to enhance peer-assessment judgements. In the main, these criteria replicate those used for selecting any workplace assessor.

Accuracy of self- or peer-assessment

Peer- and self-assessment frequently contains systematic distortions. As noted earlier, self-assessments generally are harsher than judgements by external observers.

In peer-assessment, the most common distortion occurs when judgements are overly positive. Assessors go to sometimes extraordinary lengths to avoid being critical of colleagues. Johnson (1992) studied the phenomena of politeness and compliments in his examination of peer review reports in the US. Where peer-assessment is used formatively in conjunction with self-assessment (where the contrary distortion more often appears) it provides a useful and necessary counterbalancing. In a summative context, however, such a discrepancy would create real difficulties and a process for resolving the different views would need to have been put in place prior to the assessment.

As such, self- and peer-assessment generally should not be used in forming summative judgements unless it is supported by other sources of information and judgement. In such a situation a clear policy needs to exist for the means by which the different information sources are combined to produce the overall judgement.

It has long been understood that valid and reliable assessment most likely is achieved when assessors operate within a clearly defined and consistent framework. Thus, assessment tasks must be clearly defined, as must the assessment criteria. It is not sufficient simply to provide assessors with a check list or the like. As would be true in any other case, self- and peer-assessors need to be familiar with the criteria, the context and use of the information they will produce, and must be confident in their role. Peer-assessors, in particular, need to be trained in their role. Peer assessors are not that different from any other workplace assessor. If there is a difference, it is likely to be in the frequency with which they carry out the assessor role, rather than any significant difference in the processes followed.

Multiple assessors also can improve the consistency of peer-assessments. The reliability of assessment increases as the number of independent sources of information about the assessment increases. In the case of peer-assessment this can be particularly important where individual peers might have a particular or biased view of the person being assessed. Multiple assessors allow this to be minimised or, possibly, even rejected. Of course, this has the potential to increase costs and a judgement will have to be made concerning the relative importance of accuracy versus cost.

Integration

In detailing these issues, one solution involves integrating self- and peer-assessment with information from other sources. However, it is important that, if an assessment occurs in a summative context, the question of integrating self- and peer-assessment in the final judgement must be resolved.

Peer-assessment can have a specific value in providing information on long-term performance, performance in a wide range of contexts, performance in areas which involve interaction and/or communication with peers, teamwork and the like. Self-assessment information can be particularly useful when all or some of the relevant work role is performed in a manner difficult for an external assessor to observe (e.g. a confidential counselling session). Other sources of information can then be selected which focus on other aspects of the performance criteria identified in the standards.

Appeals

It is generally recognised that formal assessment systems of any kind require some means of handling disputes (Toop, Gibb and Worsnop 1994). Peer-assessment raises the difficult issue of potentially including a wide range of work colleagues in the resolution of a dispute. The value of peer information in the assessment process must be weighted against the potential for disruption which may arise.

In self-assessment there is the potential for a learner to appeal against their own assessment (as curious as this may seem!).

Costs and resources

The cost-effectiveness of self- and peer-assessment frequently is cited by those who advocate these approaches.

There is no formal evidence to suggest that self- and peer-assessment is or is not effective and/or efficient. The cost-effectiveness of peer-assessment is yet unknown. Good assessment (of whatever kind) costs: the real issue is judging the benefits.

Anecdotal evidence suggests that self- and peer-assessment as part of an overall system could help reduce the number of instances in which trainees present themselves for assessment when they are not yet ready. If this is the case, cost savings would occur.

On the other hand, there seems no good reason to believe that self- and peer-assessment is a cheap alternative in a formal, summative assessment. Certainly, if peers are to be properly prepared for their role as assessors, then the costs of preparation—unless they are otherwise part of the education/training program—could be substantial. In any case, self- and peer-assessment is unlikely to be the sole component of formal assessment and, in many cases, will contribute only a small amount towards the final decision. In such a case its added value is likely to be small by comparison with its cost.

Future research

Clearly, the question of peer- or self-assessment in competency-based training has not received a great deal of research attention nor has it been used to any degree. Despite that, it would appear that these methods may have applications which have yet to be developed.

Research questions which this situation raises include the following:

- To what extent and in what circumstances does the use of peer- and/or self-assessment—as a learner self-selection activity—assist learning, improve the quality of summative assessment and reduce costs?
- Is current information from peers more valid and reliable than references as a means of assessing for competence in an RPL context? If so, is this a practical alternative?
- Does the use of self- and peer-assessment in an education or training program transfer to relevant situations (such as quality circles) in the workplace? In what circumstances could this be improved by other, better, training practices?

The bottom line

Competency-based approaches to education and training were introduced because they closely related to the realities of the workplace. Moreover, for many, a significant attraction was the promise of a more objective approach to assessment and certification.

Self- and peer-assessment provide a way of gaining information not readily available from more traditional approaches to assessment, particularly in workplace contexts. However, they also introduce a range of uncertainties into the assessment process which are likely to unsettle users if they become a significant part of the process of certifying competence.

In summary, self- and peer-assessment have a place in competency-based education and training. However, this is most likely to be as an adjunct to other forms of assessment and/or as a formative assessment supporting the teaching/learning environment.

Bibliography

American Consulting Engineers Council—American Society of Civil Engineers 1990, *Project peer review: Guidelines*, American Consulting Engineers Council—American Society of Civil Engineers, Washington DC.

AMA/ACHS 'Peer Review Resources Centre 1993, editorial, *Australian Clinical Review*.

Bennet 1993, 'The validity and reliability of assessments and self-assessments of work-based learning', in *Assessment and Evaluation in Higher Education*, vol. 18, no. 2, pp.83–94.

Boud, D 1980, *Self- and peer-assessment in higher and continuing professional education: An annotated bibliography*, Tertiary Education Research Centre, University of NSW, Kensington, NSW.

Boud, D, in press, 'Enhancing learning through self-assessment', Kogan Page, London.

Broadfoot, 1990, *Records of achievement: Report of the National Evaluation of Extension Work in Pilot Schemes*, HMSO, London.

Cohen, R, Flowers, R, McDonald, R & Schaasfma, H 1993, *Learning from experience counts: Recognition of prior learning in Australian universities*, commissioned report for the Credit Transfer Project Working Party on Credit Transfer and Recognition of Prior Learning, Australian Vice Chancellors Committee, University of Technology, Sydney.

Conway, R et al. 1993, 'Peer assessment of an individual's contribution to a group project', *Assessment and Evaluation in Higher Education*, vol. 18, no. 1, pp.45–56.

Dugas, J & Horrocks, R 1993, 'Using the portfolio approach to assess Australian managers: A case study', *Testing Times: A national conference on assessment for competency-based training*, National Centre for Vocational Education Research, Adelaide.

Fitzherbert, C 1993, 'Assessing clerical and administrative competencies in the workplace', paper presented at National Assessment Forum, Sydney.

Goedegbuure, L et al. 1990, *Peer review and performance indicators: Quality assurance in British and Dutch higher education*, Lemma, Utrecht, Holland.

Goldman, R 1994, 'The reliability of peer assessment: A meta-analysis', *Evaluation and the health professions*, vol. 17, no. 1, pp.3–21.

Griffin, P 1988, 'The international English language testing system: Properties of writing assessment', paper presented at the Singapore Regional English Language Centre Conference on Language Assessment, April, 1988.

Hager, P, Athanasou, J & Gonczi, A 1994, *Assessment: Technical manual*, AGPS, Canberra.

Holland, S 1993, 'Practising what we preach: Implementing RPL for staff', *Testing Times: A national conference on assessment for competency-based training*, National Centre for Vocational Education Research, Adelaide.

Hughes, I & Large, B 1993, 'Staff and peer assessment of oral communication skills', *Studies in Higher Education*, vol. 18, no. 3, pp.379–385

Johnson, D 1992, 'Compliments and politeness in peer-review tests', *Applied Linguistics*, vol.13, no. 1, pp.51–71.

Kearney, P 1992, *Collaborative assessment techniques for CBT and enterprise learning*, Artemis, Hobart, Tasmania.

Kearney, P 1993, *Collaborative learning techniques for training and education*, Artemis, Hobart, Tasmania.

Keaten, J & Richardson, M 1993, 'A field investigation of peer assessment as part of the student group grading process', paper presented at annual meeting of the Western State Communication Association, Albuquerque, New Mexico.

McCabe, R 1993, *The accountant's guide to peer and quality review*, Quorum Books, Westport, Conneticut.

National Training Board 1991

National Office Skills Formation Advisory Body 1993, *Assessor's manual: Assessing clerical-administrative competencies in the workplace*, National Office Skills Formation Advisory Body, Melbourne.

Northcote District Health Council 1988, *A resources collection: A collection of resource materials for community workers in health*, community development in health project, Northcote District Health Council, Northcote, Victoria.

Patillo, J 1984, *Quality control and peer review: a practice manual for CPAs*, Wiley, New York.

Raffe, D (ed) 1992, *Modularisation in initial vocational training: Recent developments in the European countries*, University of Technology, Sydney.

Somervell, H 1993, 'Issues in assessment, enterprise and higher education', *Assessment and Evaluation in Higher Education*, vol. 18, no. 3, pp.221–233.

Toop, L 1995, personal communication to P Griffith

Toop, L, Gibb, J & Worsnop, P 1994, *Assessment system design*, AGPS, Canberra.

Authors

Geof Hawke is senior research fellow with the Research Centre for Vocational Education and Training at the University of Technology, Sydney. He also works as a private consultant. Geof was the founding chief executive officer of the National Community Services and Health Industry Training Advisory Board and, for over 20 years, a manager and policy adviser within NSW TAFE. For much of that time, Geof headed TAFE's Assessment Research and Development Unit. His recent work includes a strategic analysis of the research and development needs of the VET system in NSW, a study of work-based learning and an examination of 'holistic' assessment.

Professor Patrick Griffin is director of the Assessment Research Centre at the Royal Melbourne Institute of Technology. He has published 16 books and more than 50 journal articles on assessment and evaluation topics that include competency, language proficiency, industrial literacy, school literacy and numeracy, profile development and portfolio assessment. His work currently focusses on interpretive frameworks for criterion referenced assessment. In this area, he has presented invited lectures on assessment matters to international organisations such as the United Nations and the World Bank. He has also addressed major professional associations, taught and conducted assessment research projects in Australia, Hong Kong, Ireland, the United States, Australia, New Zealand, Canada and Britain.

Still counting for nothing?

Bronwyn Travers

Introduction

This paper looks at the use of competency-based assessment for women from home duties or marginal labour market attachments when trying to enter or re-enter the labour market.

To discuss the assessment issues for this group we need to look briefly at some of the characteristics of women's participation in the labour force, the type of unpaid work which women perform and its effect on their assessment for entry or re-entry to the labour market. It is also important to understand the different types of women in this group in terms of a wide range of factors such as culture, language, age, levels of formal and informal training, familiarity with assessment processes, time at which they entered the labour market, paid work experience and unpaid work experience.

For the purposes of this paper, the term 'marginally attached to the labour market' refers to the Australian Bureau of Statistics definition which incorporates:

- those who want paid work but can't start work in the reference week (the week in which the ABS survey was carried out);
- those who had not actively looked for paid work but could start within four weeks;
- discouraged jobseekers.

For the majority of the women (almost 600,000, ABS 1993) in this group the most common reason for not seeking paid work was 'family related', including child care.

Women in the labour force

A lot of the recent growth in women's participation (52 per cent) in the workforce has been in part-time and casual employment. This is due to economic factors and people management strategies which have led to changes in the type of employment and the way work is contracted, as well as non-economic factors which also influence women's participation. These include women's large share of unpaid work and family responsibilities, social attitudes, traditions and religious beliefs (Rimmer and Rimmer 1994).

Women's paid employment is often interrupted—in the majority of cases this is because of child rearing (Australian Bureau of Statistics, 1994 p.141). In particular women who are marginally attached to the labour market say that home duties and child care (75 per cent) followed by education (11 per cent) are their main activity. For those who had difficulty getting a job, the most common barrier was 'balancing family responsibilities and the job'.

This pattern of interrupted paid employment means that women find it difficult to name the skills they possess both for future employment opportunities and to help with skill assessment. This also affects training and re-training opportunities which need to take into account the requirements of the large group of women who experience career breaks yet develop skills through their unpaid work.

As the study by Rimmer and Rimmer (1994) shows, repeated career breaks make it harder for women to get jobs in different areas. While this may be because some skills are not current, most of the women surveyed were out of the paid workforce for less than two years for the first career break. For later breaks the absences from the paid workforce were much shorter, with many women returning after a year or less.

We can see then that for women re-entering the paid workforce this decline of occupational status may also mean that skills developed outside the paid workforce may not be assessed as transferable to a new context. It may therefore be difficult for women to find acknowledgement for these skills from employers upon re-entry to the workforce—despite the usefulness of such skills in the workplace (see for example Cox & Leonard 1991; DEET 1992). For the women concerned there may also be difficulties in naming and therefore valuing these skills, particularly in relation to the workplace (Poynton 1993). These issues are discussed more fully in 'Skills identification and recognition' later in this paper.

Unpaid work

The Australian Bureau of Statistics describes unpaid work as 'any activity involving the production of goods and services for which someone could be paid but which has been performed for no money' (ABS 1993). It therefore includes duties and volunteer work outside the home and, not surprisingly, most of this work is carried out by women (ABS 1994; Rimmer & Rimmer 1994).

Looking at this work and estimating how much it costs has only recently occurred in Australia. In 1992 the unpaid household work performed by married women costs around $58.2 billion, with a further $5.3 billion for their volunteer and community work (ABS 1994). Looking at how much unpaid work really costs helps to show how important work is, which will benefit those who carry out these duties.

Whether the paid workforce can recognise these skills and whether the skills can be assessed to give credentials (i.e. a certificate) is an important issue for discussion in this paper. Unpaid work carried out in the home and in the community plays a major role in developing skills for many women.

Competency-based assessment

Competency-based assessment (CBA) and *recognition of prior learning* (RPL) processes have been developed and used as part of the National Training Reform Agenda. They provide assessment against clear outcomes (*criterion-referenced assessment*) to help recognise skills gained from formal training, on-the-job learning or life experience (Bloch & Thomson 1994). Ideally these processes provide a way for women to get their skills recognised including skills from areas such as unpaid work/life experience. Assessment that recognises previous learning shows that it is not important where, how

and, to some extent, when skills were learnt. In this way women, who traditionally have less formally accredited post-secondary training, more on-the-job training and interrupted employment histories, may have their skills formally recognised.

However, research shows that such processes are still subject to bias against people's gender (Lawrence, Butler, Simons 1994) and culture (Kirby 1993; OMA/DEET 1994).

Butler and Connole (1994) show how difficult it is to get all skills recognised, when both socially and industrially skills are defined by male-dominated jobs. Similarly, Phillips and Taylor (1980) argue that 'skill has increasingly been defined *against* women—skilled work is work that women don't do.'

One theory that explains why women's skills are undervalued believes that the work women do in the paid workforce has a low status because it involves the same skills used to perform unpaid domestic work. The lack of payment for that 'socially invisible' work means that it has little worth (Women's Employment Branch 1990). Research in the UK into the accreditation of skills gained from unpaid work also discusses this issue and points out that despite the opportunities accreditation provides, there is also the danger that it confirms 'women's role in the home as the basis for women entering low pay and low status jobs . . .' (Thorpe, Edwards and Hansen 1993).

As Stephens and Bertone (1995 p.21) show in their examination of the relevant literature 'gender bias in the industrial relations system . . . has led to systematic bias in the evaluation of skill with many skills and attributes applied by women in traditionally female jobs being regarded as innate rather than learned' with the result that such skills may not be considered as 'worth measuring' for the purposes of giving classifications or payment. Women from non-English-speaking backgrounds (NESB) have more disadvantages when seeking recognition for qualifications and skills which may have been gained overseas (Stephens & Bertone 1995; DEET 1992; 1993).

Kirby (1993) examines *competency-based training* (CBT), of which competency-based assessment is an important part, and its consequences for education in Aboriginal contexts and for Aboriginal self-determination. He 'draws attention to the problematic nature of CBT and postulates that, as a normative and conservative trend derived from economic rationalist ideology, CBT is in fact inimical to the ideal of Aboriginal self-determination' (p.1). Kirby states that CBT fails to 'adequately acknowledge and cater for cultural diversity' which constitutes 'an ideological position which legitimates and attempts to reproduce the culture of mainstream society within Aboriginal contexts.' CBT is seen as an example of 'mainstreaming as assimilation' where 'Aboriginal people are increasingly expected, in the name of equality to abandon specifically Aboriginal cultural practices, modes of organisation and domains in favour of 'mainstream culture and services' (p.4).

In New Zealand the recognition practices established for Maori people were developed by Maori people after extensive consultation and research within their community. These practices include the use of *attestation*, where members of a person's family and community can speak for them and provide support. The recognition process which has been developed means skills developed within Maori culture are recognised and accredited in the same way as those developed outside this culture (Benton 1991).

Bloch and Thomson (1994) show that a *best practice* approach to competency-based assessment requires that the assessment has validity. To achieve this we should look at the social consequences of the judgement being made. They go on to note that this

perspective on validity within a competency-based assessment system has not yet been acknowledged by Australian policy makers.

What are the implications of these findings for important people in the assessment process—applicants (or potential applicants for assessment), assessors and employers? How will they affect the competency-based assessment of women applicants (or potential applicants) from home duties or marginal labour market attachments who want to enter the labour market?

Australian and overseas practices

While there is an increasing amount of Australian research and resource development concerning skills identification and recognition, Hall (1994) and Thomson (1993) point out that Australia has yet to develop a body of research on competency-based training and assessment to provide a firm basis for government policy and practice. Stephens and Bertone (1995) agree with this in their literature review which look at issues facing NESB women who want training in the manufacturing industry. A recent report (DEET 1993) shows the difficulties faced by women from NESB who want their skills recognised within TAFE. The report noted that many women receive no recognition of their previous vocational or work experience—even when the paid work experience they had gained overseas related directly to their areas of study.

However, the following selected research reports and case studies drawn from Australia, New Zealand and the UK show how competency-based assessment is used for women and highlight some new practices for this group.

The findings are looked at in terms of a *best practice* (or best way of doing things) approach to assessment, with the discussion on identifying and recognising skills being useful for the application of competency-based assessment.

For the purposes of this paper, the assessment processes we have already identified will look at the validity and reliability of their use for the assessment of women seeking entry to the labour market from home duties and marginal labour market attachments.

Identifying and recognising skills

A streamlined and efficient approach to skills recognition means that individuals can identify their skills and match these to *learning outcomes* of training programs, competency standards, enterprise standards or other benchmarks. To support this process, particularly for women in this group, a number of resources have been produced for use by training organisations (see for example DEET 1992; OMA/DEET 1994; CAE/ACFEB 1995; Women's Advisers' Unit/DIR 1992).

In their report *From ummm to . . . aha! Recognising women's skills* (1991) Cox and Leonard highlighted the difficulties that women had in talking about their skills as did other research (see for example Poynton 1993; DEET 1992; Women's Advisor's Unit /DIR 1992). Robertson (1992) argues that women need to be considered as a separate group in relation to CBT systems to make up for the disadvantage they have because of their position in the training and labour markets. In terms of how much value those skills have, she gives many examples which show that the training and skills used by women

in their work are not reflected in the unskilled or semi-skilled classification given to that work. She describes skill as a 'socially and politically constructed category, . . .' and that 'those groups with most economic power are most likely to possess the label *skilled*.' However, current resource development (for example DEET 1995) supports the inclusion of all skills in competency development—the process by which the benchmarks for assessment are formed—and also encourages assessor training practices which acknowledge the experiences of women in this group.

Many people feel that the use of skills more closely associated with women than men—like interpersonal and communication skills—are not fully appreciated in feminised jobs in industries such as the clerical and service industries (Probert 1993, Windsor 1991). This is despite the fact that such skills are very important in management. Cox (1991) sees the issue from another angle and describes the current work classifications as 'men in tiny, little bitty jobs which all have got separate names, and women . . . in these broad banded jobs which have a multiplicity of skills involved in them, many of which are not named in the workplace.'

Recognition of prior learning (RPL)

Recognising prior (or previous) learning provides a way for skills and knowledge to be assessed. These skills are gained through:

- formal training (industry and education)
- work experience
- life experience.

RPL looks at the learning outcomes or competencies of these experiences not how, when or where the learning occurred. It names and assesses all the learning experiences of a person in relation to clear *vocational outcomes*—either course outcomes from curriculum or skills and competency standards from industry (Broadmeadows College of TAFE 1990).

- The report *'RPL Implications for Women'* (DEET 1992) examines issues for women in the use of RPL in training and industry and presents recommendations. Based on a national survey of RPL applicants and potential applicants as well as discussions with women in training and in industry, an important fact was discovered for the women we are talking about in this paper.

  This finding showed that RPL helps increase women's entrance to training because it recognises skills that they learnt in other jobs or at home. This is particularly the case for women who have been absent from the paid workforce and/or study for a period of time who may have learnt skills that meet learning outcomes of particular courses. The report also pointed out that women are helped by support through group processes to identify their skills as part of the RPL process.

  The RPL assessment process which was examined in this report, was consistent with the 'best practice' in competency-based assessment (i.e. the best way of doing things). However, to make sure that the assessment is valid, the report said that assessors need to be careful that women are not discriminated against in the competency standards or learning outcomes upon which the assessment is based.

A resource, *RPL—making it work for women* (DEET 1995), which was developed for RPL assessors and support people as part of the project also highlighted this need and used case studies to help assessors to assess the evidence for skills gained from unpaid work and to promote the role of support people in helping women to seek assessment.

- The 1992 project *Developing and trialling recognition of prior learning processes (RPL) Part B: Model for RPL of women gained through their life and work experience* examined the application of RPL for women seeking entry to training and the labour market of RPL against *Mayer core competencies.*

  This project explored the important role of supportive group processes for the women involved in the trial. It provided a new approach to recognising skills gained through training and education, paid work experience and life experience/unpaid work, against the Mayer competencies which were used when assessments were being made. This process involved comparing the competencies with the learning outcomes of selected accredited training programs and with industry standards. The assessment then provided the women in the pilot study with a range of options for training and employment related to their skills and linked to the *Australian Standards Framework.*

  As we said earlier, this project used the process of recognising previous learning in an 'innovative' way so that women who make up the *target group* for this paper can have their skills assessed. The project report said that using this process was expensive because of the time and resources used. However the report also believed that with more practice, the process could become more simple and cheaper to use.

Accreditation of prior learning (APL)—UK

Simosko (1991) describes *accreditation of prior learning* as a process that helps people of different backgrounds, ages and attitudes to receive formal recognition for skills and knowledge that they already have. It builds upon the idea of life-long learning and that much of this informal, uncertificated learning is the same as skills and knowledge expected of learners completing more traditional and formal learning programs.

APL against the National Vocational Qualifications (NVQs) in the UK can take place within training providers or at independent assessment centres. The National Council for Vocational Qualifications (NCVQ) has also produced resource material such as their *NVQs and unpaid work action kit* to help APL advisors and assessors in this area.

A national project was set up in 1991 to define and examine the competencies which people get through performing unpaid work in the home. The findings from this project show that this is still a problematic area. 'Recognising such competence is economically dangerous and yet economically potentially valuable, and socially just, yet widely and deeply challenging to existing power structures' (Butler, L 1993).

A practical example of how skills accreditation can work for the target group for this paper was developed by Dr Kay Sanderson from the Norwich-based APL Advisory and Consultancy Service (APLACS), who offered an innovative application of APL for women largely from home duties ('women returners'), who want jobs or training in the UK.

Again a group approach, aided by individual counselling, was used for the 'women returners' involved in the program. APLACS staff examined the unpaid work performed by this group and identified its relationship to the organisational skills found in NVQs in management. The women were assisted individually and in small groups helped by APL advisors to develop their portfolios. The portfolio was then assessed and verified.

As a result of the program, 100 women obtained Level 2 NVQs in business administration and four obtained Level 4 NVQs in management.

Dr Sanderson noted that the small group processes aided by individual counselling (as well as the APL assessment of the women's portfolios) represented a 'Rolls Royce' approach to skills assessment. It was funded under the European Union initiatives to help people to return to the workforce and she noted that it was an approach that may be difficult to keep going if the women involved had to pay the true cost of preparation for this assessment.

This highly resourced program shows how the APL process can reveal that the skills gained by 'women returners' through unpaid work are relevant and useful for other jobs. The women gained qualifications accepted throughout Australia and improved confidence and self-esteem to look at further career and study options.

The program looked beyond the technical skills involved in performing home duties and other forms of unpaid work and identified the management skills needed. In this way the organisational context in which the technical skills are used is acknowledged and pathways to a broader range of careers are established. This process also helps people to put more value on skills learnt doing unpaid work (Broadmeadows College of TAFE 1994).

Competency based-assessment (CBA) in the workplace

In this section we look at the use of competency-based assessment for women in the workplace or for the purpose of gaining employment. The areas selected reflect the pattern of women's employment, many women (65 per cent of workers) have jobs as a 'salesperson and personal service worker'(ABS 1994). These are:

- ACCESS assessment in the hospitality industry;
- competency-based assessment recruitment practices in service industry.

• ACCESS assessment in the hospitality industry

The ACCESS Skills Assessment and Recognition Program was set up in 1991 by the national industry training body, Tourism Training Australia. It was developed to support the competency standards accepted throughout Australia in the hospitality industry and to help recognise the skills of staff already working in the industry.

This program was examined as part of the Victorian State Training System Strategy for Women 1992/93 to look at women's participation in the program and their training preferences. The report on this project *Targeting training in the workplace, women, skills recognition and training* (OTFE unpublished) provides a statistical examination of how many women take part, type of jobs they have and some good ways of linking training and assessment for casual staff in the industry.

The hospitality industry has increasingly high levels of casual and part-time employment—with women making up a majority of staff employed in this way. This is a challenge for an industry which wants to upgrade skills through training, particularly where some sectors of the industry provide little, if any, training for staff (OTFE unpublished).

In its key findings the report showed that the ACCESS assessment process could be a very positive and affirming experience for women in the industry who were given a credential (certificate) that could be used throughout Australia, for the skills that they possessed.

Casual staff expressed strong interest in participating in assessment and training—even when delivered outside work hours. However, casual staff who want their skills recognised are affected by the following points:

- a tendency to undervalue their skills;
- a lack of awareness of the existence of ACCESS (even where employers were aware of the program). The majority of ACCESS assessments were conducted for full-time employees in large accommodation venues or clubs;
- a need for information about the program and support from the ACCESS registration body—Tourism Training and from employers to register for ACCESS;
- availability of practical training which would help when individual's skill levels are not up to industry's standards;
- the registration cost of $35 for candidates—where this is linked to the idea that 'there are no real benefits in it for me'.

As this report and the National ACCESS Evaluation by Tourism Training Australia indicate, the amount of assessments undertaken for the ACCESS program was considerably lower than expected. Many of the trained assessors who were surveyed were not regularly conducting assessments. This raised several issues such as the possibility of skills decay in carrying out the assessment. Assessors in industry found that the time required (off the job) for assessment and the cost this represented to the organisation were an issue for implementation.

- CBA recruitment practices in service industries

  A telephone survey of assessment practices used by three employers in the areas of catering and provision of child-care services was carried out to obtain information on assessment to help with recruitment. These were examined in the light of good practice and in terms of the needs of women seeking entry to the labour market from home duties and marginal labour force attachment.

Case study 1—catering

The national training manager of a large catering company applied a modified version of the hospitality industry's nationally-endorsed ACCESS assessment program to the selection of staff for food and beverage outlets located at a range of sports and industry venues.

Trained in assessment, the manager provided a three-day training program to supervisors including assessment skills, questioning and feedback skills, to conduct

'mini ACCESS' assessments for the purpose of recruiting staff. Many of the approximately 4 000 'on the books' or responding to advertisements for employment are mature-age women. The assessments assisted in selecting employees and using them in food and beverage outlets appropriate to their skill levels—ranging from kiosks to corporate dining rooms.

The assessment was based on the selection criteria in the job description that was provided to applicants. The process involved an interview and completion of a simulation exercise in food and beverage service in *a la carte* mode. Questioning was used during the assessment to see whether assessees knew and understood the procedures. Checklists showing the achievement and competencies were kept by the assessors and used to assist the staff who were recruited, to identify what skills needed developing. The management felt that verbal feedback and continued casual employment at the right level was very useful to the staff who gained positions or who were 'put on the books' for future events.

This process was judged by management as very successful (through reduced complaints about poor service). For those who were assessed it provided a *criterion-referenced assessment process* (where skills were tested against particular criteria) which had some of the features of the nationally-endorsed skills assessment program and was conducted according to a number of the criteria for best practice.

However, assessors noted that mature-age women applicants coming from home duties lacked confidence in both their interpersonal skills, numeracy skills ('mental arithmetic') and food and beverage service skills. This meant that they opted for employment as kiosk attendants and kitchen hands rather than work in more 'upmarket' outlets.

The time and cost constraints for this process meant that the full range of assessment methods available with ACCESS were not used. Neither were the full pre-assessment briefing and feedback sessions designed to assist help assessees. This may affect the outcomes for women applicants from home duties who may need further support to get recognition through this process for skills used in their unpaid work.

Case study 2—family day care workers

Local government councils in Melbourne employ women as home-based child-care workers, obtaining staff for this service almost always from mature-aged women who are currently raising their children, or who have had experience in this area.

The assessment process, therefore, for employment in this field asked for this experience as a pre-requisite in both case studies.

The assessment process involved the use of a variety of methods including: group information sessions, a series of one-to-one interviews, safety check of the applicant's house, observation of the applicant playing with children, the use of more than one assessor.

The councils generally offer staff development to assist family day-care workers, which can be linked to accredited vocational training programs—some of which are delivered in *fleximode* (using different types of delivery).

The thorough nature of the staff selection assessment, the high degree of independence in the job and the level of skills required to perform it (both in looking after the children and in the paperwork involved) do not seem to be reflected in the job's salary. This reinforces the points we made earlier about skills that are associated with women being undervalued.

The assessment process used to select family day-care workers satisfies many of the criteria for *best practice* in assessment where it is carried out by trained staff and where applicants are informed of all the assessment criteria. In some cases there is a group discussion at the first recruitment interview which may be helpful to potential applicants. On-going assessment of family day-care workers in their homes is one of the requirements for employment with councils.

There are large staffing costs when this assessment process is carried out, as applicants are generally assessed by at least two interviewers over two or three interviews. Referees are also contacted where possible, to provide more evidence to help with the assessment.

A recent report on family day care points out that '. . . industrial and training reform ignores the diversity within the existing workplace and certainly fails to recognise the situation of women whose employability is grounded in their domestic skill base.' While RPL can assist in this area 'it still assumes that recognition should then lead to entry to the *new world*' (WEETAG 1994).

Challenges and issues

Using competency-based assessment practices for women from home duties or marginal labour force attachments that are equitable and include everyone, presents a number of challenges. The challenges talked about in this paper are summarised as follows:

Identifying skills

The challenges include the need for this group of women to identify their skills and for assessors to become skilled in assessing them. This is particularly important in the case of skills developed when performing home duties and other forms of unpaid work. These must then be matched to competency standards which reflect the full range of skills required.

The work of Cox and Leonard (1991) has highlighted an important issue for women in gaining access to the assessment process. This is their inability to recognise, name and claim skills. Cox and Leonard point out that women require support to identify and 'own' their skills and that this may be particularly true for women from cultures where this practice is unfamiliar or not regarded as appropriate.
It is important to emphasise that for women—especially those who may lack confidence in their skills—group processes may help them to identify skills and give them support.

Gender and cultural bias in the competency standards

Several reports have looked at the issue of social and industrial gender bias and how it influences the development of competency standards against which assessment is

carried out. The work of Burton (1988) *Gender bias in job evaluation,* Windsor (1990) *Conducting gender bias free skills audits* and Women's Employment Branch (1990) *Women and award restructure—skill* looks at the skills which are often held by women but may be overlooked when jobs are evaluated and skills are carried out. This can mean that women employees are disadvantaged. They recommend ways of helping the people responsible for conducting these evaluations and audits to include all the skills involved in carrying out tasks.

The development of a valid assessment process also requires competency standards and assessment practices that include different cultures. Currently there is some research being conducted into assessment issues for NESB applicants (DEET/OMA (1994), Stephens and Bertone (1995) and Field and Mawer (in progress) and consultation for RPL processes for Aborigines and Torres Strait Islanders (National Staff Development Committee/Broadmeadows College of TAFE) which will help policy makers to take into account the needs of these groups. However, this consultation seems rather late in the implementation stage of competency-based assessment in Australia in comparison with practices overseas. How will the findings affect the current and future practices and training of assessors and the development of assessment systems?

Access to assessment

For women coming from home duties and marginal labour market attachments there are few points outside the training system to provide advice and support concerning assessment and how to prepare for it. It is also hard to find people who will conduct assessment linked to credentials that can be used in other areas. While there are developments in progress in this area in Australia, overseas models provide a much wider approach to skill assessment for this group, particularly in the case of the UK getting skills accredited that were gained through unpaid work against management curriculum. In France in 1990 a system of 'educational credit' was introduced to assist people to gain entry to further education to be 'adequately qualified to play a part in society'. This incorporates an individualised assessment of skills and personalised management of progress in a course (Jarvis 1992). This system appears to offer opportunities for training and employment for a wide range of groups—including the women discussed in this paper. The Centres de Bilan (assessment centres) encourage the identification of a wide range of skills and competencies and allow for many outcomes (DEET 1994 [2]).

In Australia, community-based organisations such as neighbourhood houses and community learning centres and health centres provide programs which are attended by many women in this group could be resourced to help to make assessment more accessible for this group.

Cost

Many people believe that all those taking part in the assessment process need to contribute to the cost of assessment. Within publicly-funded training organisations it is seen as a cost associated with the delivery of training, and some industry sectors agree with this too. However in other industries, for example in the hospitality industry, this may not be reflected in practice. In the hospitality industry it has taken longer to implement ACCESS because some enterprises do not want to accommodate the time

costs associated with assessment. Training organisations chosen by the ACCESS program to be assessment centres will also have to balance the *user-pays* principle with equity issues.

Currently costs for assessment vary widely depending on the type of assessment and where the assessment is carried out, for example as part of a training program or in industry.

While high costs and full-cost recovery are likely to present a barrier for this group, even small amounts to cover administration fees may cause difficulty. A recent study of the training needs of mature-age women in west Sydney (South West Sydney Institute of TAFE 1995) said that fees for access to training presented a barrier, and similar studies agree with this (see for example DEET 1993 for NESB women). For women who want to return to work, training programs often help with identifying skills and assessment. Making the women pay for these programs creates another obstacle to overcome.

In training organisations in Victoria RPL assessment costs range between an administration fee (commonly $20 to $50) which may cover an RPL application for 100 per cent recognition in a course, to full-cost recovery for the assessor time and resources involved.

In the UK, programs such as 'Skills recognised' provide a 'Rolls Royce' version of assessment which has helped the women involved achieve a great deal. It is unlikely that the women could meet the true costs of involvement in this program.

Assessor training

Assessors may require training to help them include the needs of this group of women into their assessment practices—particularly regarding the assessment of skills gained from unpaid work. They may also need help to conduct assessments that include and encourage people from different cultures to take part.

It is also important that they notice any discrimination against people of different gender or culture, that may appear in the competency standards against which they assess applicants.

Thomson (1993), in his paper presented at the NTCC Senior Executives' Conference, questions whether current assessor training programs (between two and five days in length) provide more than 'survival skills' in assessment. How then can assessors be given the skills to incorporate the gender and cultural issues raised in this paper?

Future directions

Research needs

We need further qualitative and quantitative analysis of competency-based assessment practices to show who is being assessed and the outcomes of the assessments for all the participants. The competency standards against which assessment is carried out also need examination in the light of the issues identified in this paper.

The development of policy for skills assessment systems which include the needs of this group require firm evidence. Particularly where the opportunity exists to recognise skills, including wide-ranging skills developed in unpaid work, against competencies in both traditional and non-traditional female jobs.

We also need Australian examples of good practice in the assessment and certification of skills for women from home duties and marginal labour market attachments who get jobs or improve their career paths. These can be used to inform assessor training programs and the use of assessment in industry and training organisations.

Resource needs

The need for promotion of skills assessment has been widely identified both in industry and training organisations and in the community. We need to show this group how important assessment is, and how and where they can take part in it. We should also encourage applicants to seek assessment as a way of developing their skills.

For the women who form the target group of this paper, showing information about the benefits of assessment may help them when looking for work after a break in their career and halting their slide in occupational status.

Providing information and support for women to continue their skills development and to document it through relevant centres, may bridge a gap that is apparent in what the services are providing at the moment.

Conclusion

We have shown in this paper that women in the paid workforce have traditionally been disadvantaged and have therefore not always been well-served by industrial changes brought about by the national training reform agenda. For women from home duties and marginal labour market attachments who want to return to work, the use of competency-based assessment gives them the opportunity to gain accreditation or certification for a broad group of skills. However, the process and delivery of the assessment needs further development so that their full range of skills are taken into account.

Specifically, there is a need for further qualitative and quantitative research to set up valid competency-based assessment practices. These should take into account the social and cultural context of the assessment and enable women in this group to achieve skills recognition which full reflects their range of skills, including management skills.

Support is also required for all stakeholders in the assessment process, to maximise the outcomes of assessment for this group. Employers need examples of cost-effective company-based assessment to help develop recruitment practices and to model the placement of women to ensure the most appropriate use of their skills. Women from this group may benefit from help with skills identification and assessment, so they can 'name' and 'claim' their skills. As well as assessment, avenues for skills development need to be available for these women, so that career breaks do not lead to a devaluation of their skills and the subsequent slide in occupational status.

Competency standards bodies, assessor training providers and government training/assessment organisations need to consider the requirements of this group when they examine both assessment processes and competency standards. In this way they can ensure that assessment provides opportunities for career advancement and skill development to provide the best possible outcomes for all involved.

Will the skills of this group of women still count for nothing? The challenge remains for government and industry organisations which contribute to the development of training and assessment systems to address this question.

Bibliography

ACT Vocational Training Authority 1992, *Department of Employment, Education and Training and ACT Vocational Training Authority: Final report on developing and trialling recognition of prior learning processes (RPL)*, ACT Vocational Training Authority, Canberra.

Australian Bureau of Statistics 1993, *Women in Australia*, Australian Bureau of Statistics, Canberra.

Australian Bureau of Statistics 1991, *Underemployed workers Australia*, Australian Bureau of Statistics, Canberra.

Australian Bureau of Statistics and Office of the Status of Women 1994, *Australian Women's year book 1994*, Australian Bureau of Statistics *and* Office of the Status of Women, Canberra.

Baxter, J, Gibson, D, Lynch-Blosse, M 1990, *Double take: The links between paid and unpaid work*, Australian Government Publishing Service, Canberra.

Benton, N 1991, *Recognition of prior learning: From hegemony to symphony: A report prepared for the Maori caucus of the New Zealand Council for Education and Training in the Social Services*, Wellington, New Zealand.

Bloch, B, Thomson, Peter 1994, *Working towards best practice in assessment. A case study approach to some issues concerning competency-based assessment in the vocational education and training sector*, National Centre for Vocational Education Research, Leabrook, South Australia.

Broadmeadows College of TAFE 1990, *Report of the Ford/TAFE articulation project*, Broadmeadows College of TAFE, Melbourne.

Broadmeadows College of TAFE Victorian RPL network 1994, Conference papers 'RPL currency of the future', Broadmeadows College of TAFE, Melbourne

Burton, C 1988, *Gender bias in job evaluation, Monograph 3*, Affirmative Action Agency, Sydney.

Butler, E, Connole H 1994, *Dancing on the coals: Re-forming the gender politics of workplace education and training, Conference paper for feminine forces: Redefining the workplace.* Centre for Gender Studies, University of South Australia.

Butler, Linda 1993 'Unpaid work in the home and accreditation', ch.5 in *Culture and processes of adult learning—a reader* (eds. Thorpe, Edwards, Hanson) Routledge, London.

Committee of Inquiry into Labour Market Programs 1993, *Report* (Kirby Report), AGPS, Canberra.

Council of Adult Education/ACFEB 1995, *Building on experience—a course in recognition of prior learning*, CAE/ACFEB, Melbourne.

Cox, E 1991, *Making the system fit people. A chance to have a say—women and employment*, House of Representatives Standing Committee on Legal and Constitutional Affairs *and* Women's Bureau of DEET, Canberra.

Cox, E, Leonard, H 1991, *From Ummm to ... aha! Recognising women's skills.* Research report for WREIP, Department Employment, Education and Training, Australian Government Printing Service, Canberra.

Department of Employment, Education and Training 1994 (1), *Education and training needs of women from non-English speaking background*, National Plan of Action for Women in TAFE, Sydney.

Department of Employment, Education and Training 1994 (2), *Assessment—system design*, AGPS, Canberra.

Department of Employment, Education and Training 1993, *Profiles of women from non-English speaking backgrounds in TAFE*, National Plan of Action for Women in TAFE, Sydney.
Department of Employment, Education and Training 1992, *RPL Implications for women*, National Plan of Action for Women in TAFE, Recognition and Assessment Centre, Broadmeadows College of TAFE, Melbourne.
Department of Employment, Education and Training 1995, *RPL—Making it work for women, A resource for RPL assessors and support staff*, National Plan of Action for Women in TAFE, Recognition and Assessment Centre, Broadmeadows College of TAFE, Melbourne.
Hall, W C 1994, *Competency-based training and assessment, Review of Research* 1, NCVER, Adelaide.
Jackson, N 1991, *Skills formation and gender relations: The politics of who knows what*, Deakin University, Geelong.
Jarvis, P 1992, *Perspectives on adult education and training in Europe*, National Institute of Adult Continuing Education, Leicester.
Lawrence, K, Butler, E & Simons M 1994, *Valuing women: New approaches to/for the Training Reform Agenda, conference paper (Reforming post-compulsory education and training: Reconciliation and reconstruction)* Centre for Gender Studies, University of South Australia.
MacEwen, S, Alison (ed) 1994, *Gender segregation and social change, men and women in changing labour markets*, Oxford University Press, Oxford.
National Centre for Vocational Education *and* TAFE New South Wales 1993, *Testing times: A national conference on assessment for competency-based training*, National Centre for Vocational Education Research Limited, South Australia.
National Council for Vocational Qualifications 1991, *NVQs and unpaid work*, NCVQ, London.
National Council for Vocational Qualifications 1988, *NCVQ information note 4: Assessment in national vocational qualifications*, NCVQ, London.
Office of Training and Further Education 1992–3, *Targeting training in the workplace, women, skills recognition and training in the hospitality industry* (unpublished).
Phillips, A & Taylor B 1980, 'Sex and skill: Notes towards feminist economics', *Feminist Review* no.6
Poynton, C 1993, 'Naming women's workplace skills' in *Pink collar blues, work, gender and technology* eds B Probert, and B Wilson, Melbourne University Press, Melbourne.
Probert, B 1993, 'Award restructuring and clerical work' in *Training Agenda*, 1:2, pp.12–15.
Recognition & Assessment Centre, Broadmeadows College of TAFE 1994, *Assessor training program resource manual*, Broadmeadows College of TAFE, Melbourne.
Rimmer, R J, Rimmer, S 1994, *More brilliant careers*, Australian Government Publishing Service, Canberra.
Robertson, R 1992, *Competency-based training systems: issues for women*, paper for the VEETAC Women's Standing Committee.
RPL—Currency of the Future 1994, *Collected conference papers*, Recognition and Assessment Centre/Victorian RPL Network, Broadmeadows College of TAFE, Melbourne.
Sanderson K 1994, *Crediting skills in unpaid work, conference notes*, APL Advisory and Consultancy Service, Norwich.
Sanderson K 1994, *RPL—Currency of the Future, conference papers*, Videoconference presentation on APL for 'women returners' in the UK, Recognition and Assessment Centre, Broadmeadows College of TAFE, Melbourne.
Simmons V 1994, *RPL—An approach to suit Aborigines and Torres Strait Islanders*, paper presented at national conference for Aboriginal and Torres Strait Islander Projects of the TAFE National Staff Development Committee, Canberra.
Simosko, S 1991, *APL: Accreditation of prior learning, A practical guide for professionals*, Kogan Page, London.
Stephens, J & Bertone, S 1995, *Manufacturing uncertainty: Non-English-speaking-background women and training* Bureau of Immigration, Multicultural and Population Research, Australian Government Printing Service, Canberra.
Thomson, Peter 1988, *The school of hard knocks. A study on the assessment of experiential learning*, TAFE National Centre for Research and Development, Adelaide.

Thomson, Peter 1993, *Current research issues in competency-based assessment*, paper presented at the NTCC Senior Executives' Conference, National Centre for Vocational Education Research, Adelaide.
Thorpe, M, Edwards, R & Hanson, A (eds) 1993, *Culture and processes of adult learning—a reader*, Routledge, London.
Tourism Training Australia 1993, *National ACCESS evaluation*, Sydney.
WEETAG 1994, *Family Day Care. Skill recognition, training packages and certification—an Australian policy*, AGPS, Canberra.
Windsor, K 1991, *Skill counts: How to conduct gender-bias free skills audits*, AGPS, Canberra.
Women's Advisers' Unit South Australia/DIR 1992, *What's in a word: Recognition of women's skills in workplace change*, Women's Advisers' Unit, Adelaide.
Women's Education and Training Co-ordination Unit 1995, *Issues concerning the training needs of mature age women in the Western Sydney region* (draft) , South Western Sydney Institute of TAFE, Sydney.
Women's Employment Branch, Department of Labour 1990, *Women and award restructuring—skill*, Department of Labour, Melbourne.
Women's Employment Branch. Department of Labour 1990, *Professional service with a smile. Recognising and tapping the skills of women workers in the hospitality industry*, Department of Labour, Melbourne.

Author

Bronwyn Travers is a project officer at the Recognition and Assessment Centre of the Kangan Institute of TAFE, Victoria. She has a BA, Dip Ed and post-graduate diploma in TESL and extensive experience in recognition of prior learning (RPL). Bronwyn has managed projects for national flexible delivery in the hospitality industry, conducted research in trade certification reviews and national RPL resources for women and undertaken senior RPL training programs nationally for industry and training providers. She has also had experience in management, development and delivery of a wide range of ESL programs in TAFE (including practical placement of ESL students in industry) and ESL program delivery overseas.

Needs of special workers

Josie Misko and John Saunders

Introduction

'In general, educational assessment has not been friendly to students from diverse cultural, linguistic and economic backgrounds' (Garcia & Pearson 1994 p.77). It still remains to be seen whether competency-based assessment will be any kinder to workers who do not fit the traditional mould.

The purpose of this paper is to examine findings of research and practice in competency-based assessment with regard to the needs of workers not normally employed for the standard working week (casual workers, part-time workers, shift-workers, seasonal workers) and workers for whom English language, literacy and numeracy are a problem.

What is competency-based assessment?

In competency-based assessment (CBA), individuals are judged according to a predetermined standard rather that in comparison with others. Industry or enterprise competency standards identify the elements of competence required, the various situations in which these must be demonstrated and the performance criteria to be satisfied. Individuals meeting requirements are judged competent and may be rewarded with increased opportunities for increased pay. From an employer's perspective CBA can help to increase the potential for workplace safety, improved productivity, satisfied customers, worker motivation and workforce flexibility.

Workplace CBA can be conducted either at the workplace or off the job. CBA occurs off the job where assessment would be too expensive (e.g. piloting planes) or where equipment would be too hazardous or inaccessible to use during production times (e.g. crane driving). Off-the-job CBA can be used to assess knowledge and understanding, as well as skills which can be demonstrated in situations which approximate on-the-job contexts (e.g. mixing drinks for bar attendants, making beds for nurses, erecting a scaffold for scaffolders).

Provided sufficient evidence can be produced, workers may also gain recognition for skills and knowledge learned through prior experience. This is done by means of a procedure known as recognition of prior learning (RPL). RPL is important to the special groups discussed in this paper because it can enable them to have skills and knowledge recognised without the need to undergo formal training and assessment.

Issues of validity and reliability

It is important that assessors of competence be capable of carrying out valid assessments because these can form the basis for decisions about job retention or increases in pay. A *valid* assessment is one which measures what it is supposed to measure. This relies on assessors knowing what constitutes competent performance of the skill, and being able to recognise competence where it occurs.

A reliable assessment will give consistent results regardless of who conducts it and where or when it is conducted. In the context of the special groups discussed in this paper, validity and reliability should not be allowed to be affected by unnecessary English language and literacy demands, nor by limitations of access to suitable assessors or facilities due to employees' peculiar working hours.

The question of what constitutes sufficient evidence will also need to be answered when assessors go about assessing competence on the job. This will be particularly important for workers who do not have full access to equipment or opportunities to practise their skills because they are not there for the full working week.

Literacy

For this paper literacy means the ability to read and write English and perform numerical calculations to levels expected in specific contexts. Some contexts require higher level skills than others. For example, insurance agents who need to read product information, write letters, calculate premiums, or access information from manuals, may require higher levels of literacy than operatives in a wine bottling and packaging plant.

For workers in general, however, literacy means being able to read job advertisements and complete job application forms, read memos, work instructions, standard operating procedures and safety signs, and access technical information from manuals. Literacy also means being able to access and read data from computer terminals, perform basic calculations required in the job and check information on pay slips.

Many adults in Australia have problems reading, writing and performing basic calculations (Wickert 1989). For example, almost a third of the sample studied by Wickert was unable to identify the gross pay for the year to date on their pay slip, 16 per cent could not correctly find an intersection on a map, 22 per cent could not locate a single piece of information in a piece of prose and 49 per cent could not consult a paint chart to determine which paint to buy.

In manufacturing and construction, literacy problems can affect safety. A survey of a number of building sites in Victoria (*English in the workplace—a research project for the Victorian building industry*) found that 20 per cent of workers would have problems reading signs like WATCH YOUR HEAD, 3 per cent would not recognise signs like DANGER, and 13 per cent would not be able to read and understand signs like HAZARDOUS CHEMICALS.

Language and communication

When we communicate we use verbal and non-verbal language to send and receive messages of intentions, expectations and beliefs. We can send or receive messages through letters, memos or notes by postal or electronic mail. We can speak to others in person, by telephone or through audio-tapes and videos. Throughout this process we can communicate pleasure, displeasure or indifference by our actions.

In CBA a basic requirement is to be able to send clear messages, and receive messages accurately. When there is an English language or literacy problem CBA becomes more difficult. A worker may be competent but not able to understand the assessor's instructions or verbalise answers to questions.

Cultural influences

The way we communicate with others is also regulated by our cultural norms. In some cultures authority is not questioned, self-praise is frowned upon and reticence is a sign of modesty. In others, the questioning of authority shows independence, appropriate self-praise reflects confidence, and being pushy is regarded as showing initiative.

In CBA it is important for assessors to take these cultural values into account (Mawer 1992, cited by Quirk 1994). For example, reticent employees may need to be prompted to divulge knowledge or explain their actions, or they may need encouragement to admit they do not understand a question or an instruction from the assessor.

The changing workplace

Workplaces in developed countries have experienced major changes in the last two decades. Industry has *retooled* to become more productive, *restructured* to become more efficient and adopted *quality assurance processes* to reduce wastage and retain competitive edge.

These changes have affected the way work is organised, products are manufactured, and the way workers interact with each other, with management and with customers. This drive for efficiency and productivity has also meant an increased reliance on more complex machinery and processes, and fewer but more-able workers to oversee the functioning of the new technology. With the exportation of many low-level skill jobs abroad there are fewer opportunities for workers with minimal skill levels. It is against this economic and industrial backdrop that the need for workers who are literate, numerate and can communicate effectively has emerged.

Increased literacy demands

These workplace changes have given workers at all levels increased responsibilities. Workers are now required to be able to read, write and communicate much more than in the past (Carnevale, Gainer and Meltzer 1990; Mickelucky 1987). Once migrants without a word of English could perform jobs in the steel, automotive and textile industries because jobs were structured so that workers only needed to know the

practices of one job well. Today these industries want their employees to become part of a new range of workplace practices which rely on workers being able to discuss work issues with their team members, act as scribes and prepare reports in English.

In service industries, such as banking and insurance, literacy demands have also increased. The drive for good customer service requires workers to know company products, and be able to understand and solve customer problems (Carnevale, Gainer & Meltzer 1990).

Increased casual and part-time work

Casual and part-time work, in the main performed by women, is becoming a permanent feature of the Australian labour landscape. The numbers of people in casual and part-time work more than doubled between 1960 and 1990 while numbers in full time work increased by less than half (ABS statistics cited by SA Women's Adviser's unit, 1991).

It is also notable that many employers are now using casual employees on a regular basis. This allows them to have the flexibility of regulating their employment numbers according to economic need, and to avoid having to provide for the benefits of a permanent workforce.

Because most employers like to depend on workers they have used in the past, the casual and part-time workforce has become more like a permanent workforce working fewer hours. Nevertheless casual and part-time workers tend to be employed in jobs requiring low-level skills or involving highly automated processes. They receive less training than permanent employees and have access to fewer career paths.

Access and equity issues

It appears that the fundamental issue regarding CBA for these workers is their right to have their current competence recognised and recorded. Workers may already be considered competent in their current jobs irrespective of employment status or language and literacy difficulties. Work history, availability and language and literacy become an issue when workers want to be considered for higher levels or a broader range of jobs which may require extra training.

Findings from research and practice

Few empirical studies deal specifically with CBA for the groups in question. Those studies which have been found to be relevant for our purposes are formal evaluations of skills recognition programs, case studies of people with a non-English speaking background (NESB) women's attitudes to training and assessment, American studies of difficulties experienced by bilingual students during traditional assessments, and studies on difficulties experienced by immigrants in having their overseas skills recognised.

In addition we have also conducted our own discussions with 39 organisations in ten major industries to obtain an up-to-date and comprehensive picture of the extent and

nature of CBA in Australian workplaces. These organisations include companies, agencies and relevant union divisions of the automotive, food production and processing, mining, energy, construction, hospitality and tourism, electrical and electronics, retail, financial services, health services, engineering, security and cleaning services industries. Discussions were also held with government agencies.

These organisations were selected to cover a broad range of industries which employed high numbers of non-full-time workers or workers with low levels of English language and literacy, and to include some of the organisations thought to be using CBA as discussed in Bloch & Thomson (1994).

The evaluation of the ACCESS skills recognition program

The tourism and hospitality industry allows workers and non-workers to register for assessment of skills with its ACCESS program for a fee. In March 1993 it evaluated this program and found that only 273 out of a possible 25 000 establishments had workplace assessors and only 360 out of 1446 registered candidates had been assessed. The three major reasons candidates gave for not being assessed were:

- waiting for assessors to find time to do the assessment;
- not having time to arrange it; and
- not knowing who could do the assessment.

When ACCESS workplace and industry assessors were asked for their opinions they agreed that employees with a low level of English language and literacy might find it hard to gain the appropriate information from written materials, complete registration forms and seek help in doing so. Assessors also suggested that marketing of the program to these workers would require materials to be written in plain English, and program information to be provided in different languages.

Although most of the candidates who had been assessed were satisfied with the process, many felt that the program could be improved by more explanation and more details of the skills and knowledge required by the units of competence. Extra study materials and videos, and more skills training and work experience, were identified as ways to help candidates better prepare for assessment. The support and recognition of ACCESS qualifications by employers were also seen to be important to the success of the program (Tourism & Hospitality Training Council, 1993).

Discussions with the South Australian Division of the Tourism and Hospitality Industry Training Council showed that between February 1994 and February 1995 there were 5466 registrations for assessment across all States and Territories. This represented more than half the number of registrations over the last four years. Registrations do not give an accurate view of the actual number of candidates because each candidate can register for a number of assessments. Nevertheless, when we take into account that in South Australia there are about 35 000 employees in the industry, of whom about two-thirds are part-time or casual workers, it is evident that very few employees and their employers are taking advantage of the program.

In cases where employers have embraced the program, however, hoteliers have found that formally assessing workers for their current skill levels, paying them accordingly and providing for further training where necessary has made good sense. Staff members do better work, incur less wastage and are better for business.

These findings can help focus government and industry attention on the need to better market CBA. If CBA fails to be adopted by industry then the benefits of planned skill development envisioned by the national training reform agenda may never be fully realised.

NESB workers in industry

Production workers in the vehicle industry can qualify for the Vehicle Industries Certificate (VIC); production workers in the metals and engineering industry can qualify for the Engineering Production Certificate (EPC). When NESB workers in each industry were asked by Stephens & Bertone (1995) how they preferred to be assessed for these certificates the great majority of those interviewed at Ford Australia (85.6 per cent) and Preslite Australia (69.5 per cent) preferred to be tested orally.

Because more than half the NESB women had found the English used in the training room too difficult to understand, the researchers surmised that this may also have been the case for the assessment process too. When the women at Preslite Australia were asked where they preferred to be tested, almost one half of the sample (45 per cent) preferred to be assessed at their work stations.

The recognition of overseas skills

NESB immigrants can apply to have their overseas qualifications assessed by the National Office for Overseas Skills Recognition. Chapman and Iredale (1990 cited by Stephens and Bertone 1995) found that 39 per cent of skilled immigrants apply to have their skills recognised and that of these less than half (42 per cent) are granted equivalent status.

Applying for skills recognition is an involved process. It requires the location of academic transcripts, and the preparation of work histories, references and evidence of performance. It may also require sitting written knowledge tests. If qualified migrants are hesitant about seeking skills recognition for fear of being rejected, or have had trouble locating and preparing appropriate information and as a result are taking lesser jobs, the cost to the Australian economy is substantial (NACSR, 1991 cited by Stephens & Bertone 1995).

Findings from American studies

More empirical studies are available when we come to language and literacy problems experienced by bilingual students in written school and college assessments. A concise overview of relevant research is provided by Garcia and Pearson (1994) who report findings that second language students were more likely to:

- experience difficulties with a timed problem-solving test because they spend most of the time working out the meaning of the question;
- be inhibited from producing correct answers, even though they know the solution, by the use of unfamiliar English words used as distractors;
- be able to demonstrate increased understanding of a mainstream language text if allowed to use their own language to answer test questions.

These findings can assist practice in CBA. For example, candidates with low levels of English language and literacy will need to be given time to understand what is expected before they are asked to demonstrate their skill. Workplace assessors should use the language of the workplace when they assess for recognition of prior learning (RPL) or workplace competencies. When NESB workers are having problems communicating with their assessors, regarding their actions on the job or their job knowledge, they should be given the choice of providing answers in their own language, through interpreters who can understand them and convey responses to assessors.

Findings from current Australian practice

Limited implementation of CBA

Our discussions with managers, supervisors, training officers, workplace assessors and organisers, industrial officers and executives in unions revealed few employers and industries were using CBA. The majority either used a formal or informal appraisal processes based on supervisors' or managers' observations of performance, or had no assessment processes in place. Employers were more likely to be using CBA where industry standards and vocations were already established. These tended to be in automotive, hospitality, electrical and wine industries.

It was evident from our interviews that the national training reform agenda and CBA had little impact on practice in the industries surveyed. The main reasons given for not implementing CBA were related to cost and time required for resource development.

Furthermore, some respondents believed that it was unlikely CBA would be comprehensively implemented in industries which depended on the contracting of work to small companies (e.g. building and construction) unless there were some form of subsidy given to employers and perhaps even compulsion to put it in place. Even then employers might resist releasing workers for assessment or training. It appeared that companies with large workforces had greater flexibility and were better placed to implement a system which required time and attention to skills recognition and training.

Limited access to CBA for casual and part-time workers

It was generally felt that there would be no major problems in arranging for the assessment of permanent part-time workers, shift workers or workers who were placed on regular rosters because they could be assessed like any other worker. This was especially the case when supervisors and leading hands acted as workplace assessors.

Conditions for other casual workers were different. John Martins, a large South Australian department store chain, whose workforce comprised 70 per cent casual and part-time workers, only called on casual workers during busy times. This made on-the-job CBA inconvenient. The problem was compounded by the fact that many of these casuals work only a small number of hours each week (e.g. Saturday mornings for university students) and have other commitments which hamper their ability to attend work for assessment at other times.

In other industries casual workers may also be denied assessment because:

- employers may not want to assess workers who are not likely to remain with them for long;
- they are brought in during peak periods when there is little time for assessment;
- they may not be aware of the options available to them; and
- they may not remain with employers long enough to be assessed.

In addition it may be the case that, like some other workers, casual workers avoid being assessed because they are apprehensive and unsure of what it could mean for them or because domestic responsibilities prevent them from being involved. If CBA is to become accessible for these groups then these needs will have to be addressed.

Increased emphasis on literacy and language training as preparation for CBA

Our discussions revealed that by far the greatest amount of industry attention has been directed to workplace English language and literacy issues. This is not surprising when companies like Kellogg and ACTIL report almost 50 different nationalities on their payrolls.

The extent to which these problems with English language and literacy occur across the workforce varies between industries. For example, oil refineries like Shell which employ large numbers of qualified tradespersons and professionals, and attract applicants of relatively high ability, report few difficulties. However, fruit processing companies, and cleaning and catering sections of major hospitals and casinos, report larger proportions of workers with lower language and literacy levels.

The importance of adequate English language, literacy and communication skills for occupational health and safety was highlighted by the majority of people we spoke to. Employees must be able to read signs, danger tags, safety notices, site plans and product labels. They must also be able to understand and use safety related terminology and in some industries, communicate clearly by phone or two-way radio. In addition, BHP–Whyalla, GMH and ACTIL emphasised increased reliance on literacy and communication skills for quality assurance and improvement processes.

Benefits of CBA

CBA allows a worker to be assessed on the job (by observation of work practices, by inspection of the work produced and by oral questioning of the worker on the job, or at least, in the workplace). This form of assessment offers special benefits to the groups of workers who are the subject of this paper.

Casual and part-time workers may benefit from being able to be assessed in the course of their normal work, without the need to find additional time for assessment. Workers with low levels of English language and literacy may also benefit: firstly, from reduced dependence on English language and written communication as a means of demonstrating their ability and knowledge of the job: secondly, because their assessment is in familiar surroundings, using familiar equipment, and based on their every-day work practices. This last factor is especially important to persons who, if they are not able to express themselves well in English, may find it difficult to ask for a particular tool they need or explain a process on unfamiliar equipment.

CBA for these special groups should incorporate these particular features to the maximum degree practicable so as to gain full advantage of the benefits this form of assessment offers.

The recognition of prior learning (RPL)

Recognition of prior learning (RPL) is especially important for casual workers and those with English language and literacy difficulties. For casual and part-time workers it is one way to obtain recognition for skills accumulated over the years between companies. For those with low levels of English language and literacy it is a way to obtain formal vocational qualifications without the need for written tests. Proving that their skills are current may be problematic for those workers whose employment histories have been fragmented and inconsistent. Explaining past histories may be difficult for those with low language and literacy levels.

An innovative method for implementing RPL has been initiated at the Mitsubishi plant in Adelaide. Workplace assessors have been trained to conduct RPL alongside TAFE personnel. An RPL candidate can bring a shop steward along to the assessment as an advocate. Because the shop steward has worked with the candidate he/she can prompt the worker to recall instances when they have actually performed a skill successfully. When there is a language difficulty with English, interpreters or translators are also used.

At the time of writing the assessors had tested 400 workers with an average of 20 credit hours per worker being obtained through this RPL process. Initially shop stewards regularly attended the assessments but as the program has progressed they have become less apparent. Many of the NESB workers have expressed a desire to be tested in English with only a few requesting a translator.

For many NESB and ESB workers with low levels of English language and literacy, RPL may seem time-consuming and frustrating. In general they will require assistance with the completion of forms and some may need encouragement to even consider being assessed. They will require assistance with the preparation of RPL evidence and support in the discussion and presentation of this evidence to assessors.

Strategies for CBA: Current practice

If CBA is to become a reality for all workers in the Australian workplace then special strategies need to be implemented to meet the needs of casual and part-time workers and workers with low language and literacy. In this section we focus on strategies that have been highlighted by findings of research and practice.

Promoting an awareness of CBA

If we are to provide for appropriate assessment of all workers the commitment of employers is a fundamental requirement. This can be done by employers fostering an organisational culture in which CBA is an accepted part of normal practice. Employers will need to signal their commitment to this culture by providing the resources and time for workers to improve wages and career options through such a system.

To promote the culture, workers, unions and employers will need to meet to decide how the new cultural norms related to CBA should be marketed, programs established and rewards decided. Special efforts will also need to be made to hold these meetings at times when casual and part-time workers can attend and in ways in which workers with low levels of English language and literacy can be involved. This means meeting during work hours so that family and other external responsibilities can be fulfilled. It also means ensuring that communication with workers is straightforward and easy to understand.

Regular meetings such as these can be effective venues for canvassing input about how best to implement CBA, and explain company expectations and procedures.

CBA for casual and part-time workers

Offering flexible delivery

Flexible delivery is a client-centred approach to learning which offers workers the opportunity to learn and be assessed in ways which are tailored to their individual needs. For example, casual or part-time workers might prefer self-paced learning of knowledge in their own home, learning of skills in a training centre, followed by assessment of skills and knowledge in the training centre. Workers with English language difficulties, on the other hand, might prefer to have pre-existing skills recognised through RPL and then undertake all learning in the workplace with the assistance of English language trainers, this to be followed by assessment of knowledge in the workplace and assessment of skills on the job.

Use of the buddy system

Although primarily focussed on training, the buddy system also has relevance to assessment, particularly for workers with low levels of English language and literacy. In the buddy system, two workers are paired on the job, either so one experienced worker can provide support and assistance to the other, of so two workers can share their knowledge and experience as they learn together. If a worker with low level language or literacy is paired with an experienced worker with good English language skills, the experienced worker can not only help the low level worker master the language requirements of the job, but could also help the worker prepare for assessment by means of informal practice assessments of both knowledge and skills.

Setting up external assessment centres

Part-time and casual workers in many industries are generally overlooked when it comes to assessment or training. Where CBA is in place, however, regular casual and part-time workers may be assessed by the normal procedures. Where there is no CBA in place, or where employers are loath to provide assessments, external assessment arrangements like the ones set up by the tourism and hospitality industry may be an option. Here anyone can register with the ACCESS program. If they are not attached to a workplace or unable to be assessed at their workplace they may approach the Industry Training Council to be assessed. The council then arranges for assessments to be conducted by registered workplace assessors at TAFE colleges, host companies or private providers. Such arrangements can also be used by trainees who have no workplace, or unemployed people wanting to have their skills assessed.

Setting up easy-to-administer recording systems

The accurate recording of competencies is integral to the success of CBA. This means that workers and their supervisors or managers need to ensure that the appropriate information is recorded as objectively as possible. The timely and accurate recording of competencies is especially important for casual and seasonal workers.

One union organiser in an industry which has as yet to introduce CBA suggested that in his industry casual and part-time workers could be issued with a hard covered logbook which would contain a stamped record of competencies achieved. The logbook could then be transported between jobs with competencies achieved signed off by each assessor. A central databank held by the relevant industry training council or company personnel departments would maintain records of this information. Of course such a system will depend on the diligence of those who have to maintain such logbooks.

The Electrical and Electronics Industry Training Board (EEITB) has experimented with a number of recording methods. Apprentice progress at work has been generally recorded daily on cards. Information on these cards has been scanned into an ASCI computer file and downloaded onto a database. This system has worked in companies like ICI at Botany and when employers have been diligent in filling out these cards.

The EEITB in Victoria has recently commenced piloting the use of a hand-held electronic 'data logger' not unlike a small calculator. This data logger remains with the apprentice. Here the relevant information is punched in by the apprentice and manager. The system is easy to operate and it only takes about 20 seconds to record a day's work. A time clock is activated to record the date and time of each entry. At the end of a certain period (4 weeks or 8 weeks) the apprentice downloads this information via a modem into a personal computer. Such a system could very easily be used by casual or seasonal workers.

Providing access to records

Casual and part-time workers will need to have ready access to records of their competencies. Central databanks maintained by the relevant ITAB, personnel department, union, or assessment centre can also be charged with the issuing of transcripts for a fee. Such a system is also used by academic institutions.

Setting up special assessment arrangements

Casuals are generally employed to deal with increased levels of production. Where they are employed on a regular casual basis (e.g. seasonal workers in the fruit and wine industries) special arrangements could be made to implement CBA outside the peak period. For example, seasonal workers in fruit and wine industries could be brought in to receive training before the harvesting or processing starts. They could be kept on for a few days after the peak period is over to complete assessments. Such a scheme may also require government subsidy for it is unlikely that employers would have the funds to allocate to these extra commitments. This principle assumes special significance for workers with language and literacy difficulties.

CBA for workers with low levels of English language and literacy

Matching language and literacy requirements to the job

The language and literacy levels used in a competency assessment should always match those required by a worker to perform the job. For example, it would be unjust to give cleaners a complex written knowledge or theory test when all they need to be able to do is to recognise cleaning materials and safety signs. Conversely, workers such as electricians who need to be able to read industry standards and technical manuals should be tested on their ability to perform these functions.

Using plain English

When there is an increased reliance on print information, this will pose definite problems for employees with low English language and literacy levels. These problems can be partly alleviated for these employees by presenting information in plain English.

At Kellogg, specialists are employed to vet all training materials and assessment questions for difficult language. At Mitsubishi, language and literacy experts from Onkaparinga TAFE have prepared easy-to-understand training materials for the Vehicle Industry Certificate. The emphasis in both cases is on plain English for easy communication.

Using workplace literacy programs

Many companies have used voluntary workplace literacy programs to assist workers with language and literacy problems (ACTIL, Stanwell Power, Port Authority of Geelong, Mitsubishi, GMH). Programs which integrate the English language and literacy skills of the classroom with those required on the job have worked well. Such findings have also been supported by Sefton (1994, cited by Quirk 1994) and Courtenay (1994, cited by Quirk 1994). Examples of programs which provide language/literacy training in the workplace are: the Workplace English Language and Literacy (WELL) and the Food Industry Language and Literacy Initiative Program (FILIP), both of them federally funded.

At Mitsubishi, a workplace literacy program run by Onkaparinga TAFE, has been established to assist workers who wish to prepare for and be assessed for the knowledge modules of the Vehicle Industries Certificate (VIC). This program helps English speaking background (ESB) workers. NESB workers with low levels of language and literacy work through the modules of the VIC in workplace classrooms. This program has been called the *workplace communication program* to avoid any bad associations workers may have with terms like *literacy*.

The language and literacy instructors at Mitsubishi have found that such ESB and NESB groups need to be treated in different ways. They have placed ESB workers and NESB workers who speak some English in the same groups. NESB workers with very little or no English have been placed in their own groups. Both groups work on the same VIC modules. The modules have been written in plain English and the language that is used is the language of the workplace. That is, workers are asked to solve problems that are relevant to their work.

It is sometimes the case that workers with low levels of English language and literacy are apprehensive of classroom learning and written tests. In this program instructors work with students in small groups to increase their confidence and self-esteem. Increased self-esteem has enabled these workers to achieve success in VIC assessments.

Although it is generally expected that the NESB group will take about 50 per cent more time to complete the VIC preparation and assessment program than the ESB group, some NESB group members have learnt so quickly that the setting up of a fast-tracking system is being considered.

Using job performance aids

Job performance aids like step-by-step instructions, flow charts and computerised guidance have been used to overcome problems created by low levels of English language and literacy. Mickelucky (1987) reports studies by Elliott & Joyce (1971), Kamman (1975), and Swezey (1977) which have shown that the use of job performance aids can improve performance. High school students with 12 hours training in the use of job performance aids to diagnose electrical problems could detect faults at comparable rates to electricians with seven years experience. Comprehension was increased and reading errors and reading times were reduced by use of a flow chart. Workers with high ability, using job performance aids, did better than those using traditional print materials.

Although the use of job performance aids can assist all workers to perform their work competently, it can be particularly beneficial to persons with low English language and literacy levels. For example, pictorial and colour-coded materials have been found especially useful in helping such workers at the Adelaide Casino (cleaning section), Royal Adelaide Hospital (catering section) and Earl of Zetland hotel to understand instructions and procedures. However, reliance on these aids may mean that the particular competence achieved is limited to other jobs in which such aids can be similarly provided.

Using oral questioning

When workers' skills prevent them from fully understanding instructions and completing written tests, a number of companies have used oral questioning (BHP—Port Hedland, Shell, Stanwell Power, Kellogg, GMH, EPU). At BHP—Port Hedland the oral questioning was initially conducted by an assessor who was unfamiliar with the job. This was to reduce the possibility of the assessor inadvertently giving the candidate clues to possible answers that the assessor would enter on the assessment form. The form was subsequently marked by another assessor knowledgeable in the job, who further queried the candidate on unclear answers.

At Kellogg an assessor could rephrase test questions so that they could be better understood. For example, prompts such as 'What would you do if . . . happens? allowed the assessor to explore underpinning knowledge and understanding.

Using interpreters

Interpreters can help those with little English to understand instructions and respond to questions. In the past workers were often placed in groups of similar cultural and language backgrounds to help ease of communication. In some companies this is still the case, and workmates who understand more English often act as interpreters.

The effectiveness of using interpreters in CBA, however, is reduced if they lack the technical knowledge to appropriately relay messages between assessors and candidates. The training of interpreters in the basic language of the workplace and in assessment procedures are ways to deal with this problem. Another way is to make sure that workers from a variety of language groups and with relevant technical skills are trained in interpreting for assessment purposes.

Upgrading selection processes

The extent of language and literacy problems in the workplace can be regulated by selection processes. In view of increased workplace demands for literacy and language skills, many industries are looking at making these skills essential criteria for job performance and therefore for selection. Companies and industries with high selection requirements also reported fewer problems with low language and literacy.

Such selection practices may in fact work to discriminate against NESB workers with low language skills moving to new jobs in the event of retrenchment (Stephens & Bertone 1995). Bridging and preparatory education programs could contribute significantly to prevention of such discrimination. For example, funded programs to help employers implement language and literacy training in the workplace such as Workplace English Language and Literacy (WELL) and Food Industry Language and Literacy Initiative Program (FILIP) can encourage employers to look more favourably on applicants with low levels in these skills. Programs like the Special Intervention Program (SIP), which provides language/literacy training for the potentially long-term unemployed, can also help alleviate the problem by enabling applicants to meet the standards set by employers.

Implementing intensive coaching and retraining programs

Sometimes workers may continue to make the same mistakes because they have never learned correct applications of rules in numerical calculations, spelling and composition. At ETSA, a skills enhancement centre was set up to give two weeks intensive coaching to workers who were experiencing difficulties with assessments. A specialist literacy/numeracy trainer was hired to coach workers in work-related numeracy. The program showed that intensive coaching can ensure that workers focus on certain skills and gain sufficient practice to apply these skills once they leave training.

There is no reason why similar programs could not be mounted in many enterprises. A *brush up your skills* program could target basic and advanced literacy skills at different times during the year.

Encouraging participation

One of the greatest problems in conducting literacy and language programs which will prepare workers for assessment is to get those in need to participate. At GMH older workers tend not to want to do the Vehicle Industries Certificate. Some may feel embarrassed by low literacy levels and therefore not select themselves for assessment even though they may be highly competent in their jobs. Others may just fear exposing themselves to failure. Still others may feel that they don't really need it at this stage of their lives.

CBA at BRL–Hardy is not compulsory. Because many workers may lack confidence in their ability to pass the assessment the supervisor, who is also the workplace trainer and assessor, takes an active part in encouraging workers to consider applying for assessment and training.

One way to encourage such workers to think about applying for assessment or training is to persuade them to assess their own literacy and language skills and compare these with the skills they require to be considered competent. Ideally such reflection should make workers aware of the skills they need to acquire and encourage them either to apply for assessment or attend relevant training.

It must be remembered, however, that higher levels of language and literacy are required for training than are necessary for effective job performance (Courtenay, cited by Quirk 1994). It is often because of this that workers do not focus on their language and literacy skills until they are required for the next job.

In addition the increasing reliance on self-paced learning in vocational education may also hinder the success of workers whose language and literacy skills are deficient. Integrating self-paced learning with assistance from specialist trainers who can guide participants through the learning materials may also increase participation.

Getting people to participate in CBA requires special marketing. Workers need to see real benefits in having their current skills assessed. Only then will they willingly become part of the assessment process.

Providing extra time for skill development and assessment

Another way to help groups with low levels of English language and literacy cope with tasks requiring these abilities is to allow extra time for preparation for assessment. This strategy has been adopted at Ford, Mitsubishi, Stanwell Power and BHP-Port Hedland. When a group of steel-fixing trainees were experiencing considerable difficulties with TAFE off-the-job training in plan reading, the manager of Stanwell Power withdrew the trainees from the college. The trainees were more successful when they were given extra time to develop plan-reading skills in workplace training. At Mitsubishi, groups of workers with low levels of English language and literacy are also given extra time to complete the knowledge modules of the VIC.

At BHP—Port Hedland extra time is also given for the completion of work tasks under assessment conditions. This means that workers are allowed more time to do the assessment task than they would normally be allowed on the job.

Providing specialised resources

Specialised resources will be necessary to allow for the participation in CBA of workers with limited English language and literacy. At Stanwell Power, specialised resources were produced for use with trainees with limited reading ability. Manuals were written in plain English, and text was accompanied by pictorial information. Specialised materials and videos were also produced to go with specific modules of work. These course materials have also been used on construction sites throughout Australia.

At Kellogg, English specialists were hired to make certain that manuals and training and assessment materials and procedures were checked for language which might be difficult to understand.

Industrial relations

Discussions held with organisers and executives from a number of unions (Miscellaneous Workers Union, Liquor and Hospitality Union, Construction, Mining, Forestry and Energy Union, the Automotive Food, Metals and Engineering union, and the Electronics, Electrical, Plumbing and Allied Workers Union) revealed that these unions were generally supportive of the concept of assessment and training for their workers. They saw it as a positive way for members to be recognised and rewarded for their current skill levels and for developing new skills. It was also evident that in the majority of cases CBA had either not been introduced or, if it had, was not widespread.

Unions in industries yet to introduce CBA were concerned that employers may not support it because it would mean increased costs. This was especially the case for contractors, and other small enterprises. Concerns were also raised about the costs of implementation, selection and training of assessors, selection of candidates and setting up of grievance procedures. In some cases debate still centred on whether to pay for acquired skills or skills used on the job. It was also generally felt that unless there were some form of subsidy or compulsion for employers to be involved in CBA it would not get off the ground.

After the assessment

In common with other workers, those from the special groups focussed on in this paper may benefit in a variety of ways from an appropriate system of competency-based assessment which enables recognition of their true level of competence. Such recognition by the industry and employer is likely to contribute to increased job satisfaction, provide a sense of achievement, improve confidence and self-esteem, and encourage the worker to set career goals and consider further training.

From the employers' perspective, accurate assessment of these workers' competencies by means of more appropriate assessment can contribute to a more flexible multi-skilled workforce, contribute to productivity by more accurately matching workers' skills to job requirements and ensure the workers are aware of safety requirements. In addition, the increased confidence and self-esteem of employees can encourage them to take greater interest in their job and consider further career training in the light of company goals.

Provided the assessment is seen as fair, and provided useful feedback is given and appropriate training programs are available (e.g. language/literacy) there can be positive outcomes even for unsuccessful assessment candidates. The provision of an effective post-assessment support system can encourage them to continue with their pursuit of competence.

Costs of CBA

Costs specifically associated with the CBA for part-time or casual workers will concern the setting up of registration processes and assessment centres which will be available when these people do not have access to workplace assessors. Where CBA is in place, the cost would relate to setting up a process for record keeping and the establishing of

a central databank, similar to that used for recording superannuation benefits and workers compensation for workers in the construction and building industries.

Where external assessment arrangements are in place it would be feasible to expect the worker to pay something towards registering for assessment. However, the fee should not be too high, especially as casual and part-time workers may have limited funds. Where a log-book or an electronic device is used it would also be feasible for the worker to pay something towards the cost of the system.

Costs associated with assessing groups with low levels of English language and literacy concern the promotion of the scheme, payment for interpreters if required, for training and preparing documents for assessment. The translation of documents, forms and modules into plain English would also need to be included.

Balanced against the costs of providing assessment which is appropriate, and therefore fair, to workers from the special groups considered in this paper, must be the costs of not providing such assessment. If assessment (and training) systems do not take into account the special needs of these groups, individual employers, industry, and the nation, will not reap full benefit from the valuable resource these groups represent. Bean (1990) for example, points out the there is an enormous wealth of skill, talent and motivation locked behind barriers of language, literacy and poor organisation communication systems.

Persons with the job skills and the language and literacy skills needed to perform a job should not be barred from qualifying for the job because they are unable to cope with the language and literacy requirements of the assessment system. Similarly, casual and part-time employees should not be barred from qualifying because they are unable to gain access to assessment on the job.

Apart from the cost of not utilising an important resource, failure to provide appropriate assessment for these groups can also result in additional costs attributable to high failure rates, repeated assessments, high attrition rates from training programs and the provision of unnecessary training to meet assessment requirements rather than job requirements.

There is also a social cost: in this era of rapid technological change and industry restructuring, workers who remain classified as unskilled or low-skilled because they are unable to access appropriate assessment are likely to become the unemployed of tomorrow. On the other hand, as noted earlier in this paper (evaluation of ACCESS program), recognition of workers' current skill levels through assessment can result in better workers and improved productivity.

Conclusions

Our research shows that CBA is not widespread in Australian workplaces and that more work needs to be done to market the concept to employers and employees alike. It is also evident that many employers, particularly the smaller ones may, at the very least, need a considerable amount of encouragement and support if they are to be induced to implement it.

Because on-the-job CBA requires special time to be set aside for assessment, the work schedules and outside commitments of casual and part-time workers may prevent them from becoming aware of and involved in CBA. In addition, employers may not wish to allocate time and resources for the assessment of workers whom they do not regard as part of their permanent workforce.

On-the-job CBA relies on the demonstration of competence. This should present less of a barrier to workers with language and literacy difficulties than traditional forms of assessment. However, unless special precautions are taken, workers with these difficulties are still likely to suffer some disadvantage. To provide maximum benefit for these employees, English language and literacy training, and support schemes which aim also to boost the confidence and self-esteem of workers, need to be implemented along with CBA.

Implications for further research

There is a real need to document existing practice in those industries which have their vocation and CBA in place with regard to:

- the numbers of non-full-time workers and workers with low levels of English language and literacy presenting for assessment;
- the assessment experiences of these workers and their supervisors;
- the effects of successful CBA on workers' confidence, self-esteem, workplace participation, job satisfaction and career planning;
- the outcomes of unsuccessful CBA, particularly with regard to subsequent training and perception of career goals;
- the impact of language and literacy training programs on performance in CBA;
- the impact of successful CBA practices on productivity; and
- the effectiveness of recording and reporting arrangements.

As well as surveying employees it is also important to learn from the problems and observations of workplace assessors. This information can be used to develop recommendations for more valid and reliable assessment instruments, comprehensive and relevant assessor training and effective testing, recording and reporting processes.

References

Bean, R 1990, 'The problem of employee communication: Work illiteracy costs more than $3bn a year', *Weekend Australian* 23–24 June.

Bloch, B & Thomson, P 1994, *Working towards best practice in assessment*, NCVER, Adelaide.

Carnevale, A P, Gainer, L J & Meltzer, A S 1990, *Workplace basics: The essential skills employers want*, Jossey-Bass, San Francisco.

Garcia, G E & Pearson P D 1994, in *Review of research in education*, ed. L Darling-Hammond, American Educational Research Association, Washington.

Mickelucky, L 1990, 'Basic skills impediments to communication between management and hourly paid employees', *Management Communication Quarterly*, vol.3, no.4, pp.452–473.

Quirk, R, 1994, *Language and literacy issues in competency-based assessment*, Assessment Centre for Vocational Education, St Leonards, NSW.

SA Women's Adviser's Unit 1991, *Same time next week: The extent and nature of part-time work in South Australia*, South Australian Department of Labour, Adelaide.

Stephens, J & Bertone, S 1995, *Manufacturing uncertainty*, Australian Government Printing Service, Canberra.
Taylor, M C, Lewe, G R, Draper, J A (eds) 1991, *Basic skills for the workplace*, Culture Concepts Incorporated, Toronto.
Tourism and Hospitality Industry Training Council 1993, *Access evaluation*, Tourism Training Australia, Sydney.
Wickert, R 1989, *No single measure: A survey of Australian adult literacy*, DEET, Canberra.

Authors

John Saunders has an extensive background in TAFE, spending the first ten years as a lecturer in automotive engineering and human relations followed by nineteen years in student services—initially as a counsellor to students and staff and subsequently as a manager of student services. He is a qualified psychologist and has for many years been involved in research focussing on the development and application of assessments in a variety of fields. Having recently retired from TAFE, he now works privately as a psychologist and a consultant in vocational education.

Dr Josie Misko is a research fellow at the National Centre for Vocational Education Research Ltd with particular interests in the national training agenda and how it will affect training and assessment in educational institutions and in the workplace. She also has research interests in the transfer of knowledge, the development of expertise, and information processing. In addition, Josie has had considerable experience in developing and delivering training services for industry and spent five years in an industry training and development role. She has formal accreditation to conduct quality service and improvement training, supervisory and management skills training and personality type indicator training. She has written in the areas of transfer of knowledge, learning styles and flexible delivery of vocational education and training and has published in these three areas. She has recently completed an empirical study of the relationship between practical performance and knowledge and understanding, and another on the learning preferences of apprentices and trainees in five different occupational areas.

Competency-based assessment in remote locations

Russell Docking

Introduction

There is a growing trend to use competencies in the planning and delivery of education, vocational education and industrial training (Gonczi 1994 p.27). Competencies are also being used at work to support such processes as recruitment and selection, work redesign and performance management. A critical need for all these developments is the use of high quality competency-based assessment.

Generally, we can see that the infrastructure, policies and developments have been directed towards the use of competency-based approaches at work and during work-related training. This has meant that attention is being given to large industrial or population centres, but has resulted in the needs of people in *remote locations* being mostly ignored.

In this paper, we look at ways of managing competency-based assessment in remote locations. At first sight, a competency-based approach to training and assessment would seem to deal with most of the disadvantages experienced by those learning at a distance. In particular, it could allow for *self-paced learning* (learning at an individual's own pace), using a variety of resources and recognising competencies learned through work and life experiences. The reality might be a little less rosy, as competency-based approaches to assessment need complicated, difficult structures that may be hard to keep going at a distance.

We need to find out the potential and the problems of competency-based assessment in remote locations and to suggest ways of helping distance learners to benefit from the competency reforms.

The paper begins with an outline of the characteristics of remote locations and a brief look at the relative merits of using traditional assessment and competency-based assessment in these places. Some Australian and overseas practices are then looked at, leading to a discussion of issues and challenges for remote assessment. The paper finishes with a discussion of *cost-effective* approaches (or approaches that will use money wisely), future directions and suggestions for further development.

Remote locations and assessment

What are 'remote locations'?

Remote locations are places that are geographically far away from major population centres. They may be small communities such as mining towns, in remote tourist locations or on remote cattle stations.

People in remote locations have limited access to information and support; communication, equipment and technology; training or work experience; and experts and opportunities for the use and practice of competencies. They are isolated from others who might share their interests and ambitions, from those who might help them achieve those ambitions, and from those who might be able to assess their achievements.

Most of our education, training and assessment systems assume personal contact. They are structured for 'face-to-face' teaching and learning, so those in large centres find it difficult to understand the special needs of those in isolation.

What is 'competency-based assessment'?

How is competency-based assessment different from traditional approaches to assessment and what does this mean for remote learners?

Competency-based assessment has two key features:

- the *specification* (or naming) of clearly defined standards, (Hager et al. 1994, p.11); and
- the separation of *evidence gathering* from *judgement*. This means that getting *evidence* of a person's ability or competence should be a separate process from actually *judging* how competent they are. (Hager et al. 1994, p.5).

Setting standards of achievement

The standards are usually given as *units* of competency (large chunks of competence, such as a task or a training module), *elements* of competency (smaller components of competence that are needed to achieve a unit) and *performance criteria* (the evidence needed to show if a person had achieved an element).

The clear specification of absolute standards is different from how traditional standards are set using *norms*, such as averages. The use of norms to set standards needs a large group of candidates who have been given the same training and who have been given assessments that are *standardised* (or the same) and delivered in a *controlled* or supervised environment. (Kleinschafer 1992). Neither of these requirements is easy or desirable to accomplish in remote locations.

The specification of standards using competencies means that achievement of a competency can be recognised no matter how or where it was learned, assessment of that competency can happen in almost any location. It also means that gathering evidence about a candidate can be a separate process from forming a judgement about them.

Gathering evidence of achievement

With traditional assessment, a sample of knowledge and skills is assessed and the overall performance of the learner is judged from that. If they get 50 per cent of the sample right, we would assume they would get 50 per cent of everything right if we tested it all.

Traditional assessment also places an emphasis on how well a person does, in comparison to the whole group which is assessed. The sampling process makes this method attractive because it is quick and cheap, but it is also its weakness because it gives no detailed information about how competent a person is. The sample evidence of achievement gathered has little real worth other than as a way of guessing about overall performance such as a grade or mark. This makes the evidence gathering and the judgement process inseparable and encourages assessment made from impressions only. (Docking 1991).

Improving these tests usually involves using *multiple* (or many) assessments and/or multiple expert judges, often as assessment panels. The need for the assessment to be conducted by one or more experts severely restricts the opportunity for learners in remote locations, because the assessor must be there in person to gather the evidence and make the judgement—or impose inflexible and often invalid standardised tests from a distance.

In contrast, competency-based assessment covers everything and measures what a person can or cannot do. By separating evidence gathering from judgement, competency-based assessment lets those in remote locations play a more active role in their own assessment. Wherever they are located they would be able to collect evidence of competence (and get it authenticated where necessary) which can later be interpreted by judges somewhere else.

The requirements of competency-based assessment

For such benefits to be available for remote learners, a competency-based assessment system will need to be well designed. In particular the systems should contain the following:

- Clearly defined competency standards (units of competency, elements of competency and their performance criteria). These may be developed through looking at work practices in industry or through reviewing training objectives.
- Clearly named procedures for gathering evidence about achievement of the units, elements and performance criteria. These might include documents in a portfolio, structured observation by an assessor or formal assessments. They may, in some circumstances, need specialist assessors or on-the-job assessments.
- A competency-based record-keeping system. It is important to record everything that a person can do, not just the number of marks or the grades they have earned. As the competency record gradually builds up this information needs to be managed securely.
- Clearly defined decision-making rules to help interpret the records of evidence. These rules may show which competencies are needed for certain occupations, for entry into training, for awards or for a qualification.
- A *quality assurance* system to make sure that all of the above components are credible and relevant.

Any competency-based assessment system must meet the requirements of *validity* (it measures only what it is supposed to measure), *reliability* (such measurements produce consistent results), *flexibility* and *fairness*.

Once established, such a competency-based assessment system can provide a basis for the *recognition of prior learning* (or recognising what has been learned previously), measuring education and training achievement or assessing achievements at work.

Australian and overseas practices

> *Non-metropolitan Australia is not well served by tertiary education. Governments have tended to treat the whole of Australia in the same way, frequently ignoring the special problems faced by rural communities.*
>
> (Hall 1990)

Australia is a highly urbanised country with most of its population in major city centres; yet Australians are very aware of its large inland spaces and have a strong sense of identity with those who live there. It is surprising, therefore, to see how little reference has been made to the special assessment needs of learners in remote locations. For example, not one of 63 papers presented at a forum on quality distance education (Atkinson et al. 1991) looks at assessment issues and only one paper makes even a vague reference to the topic.

Looking at research papers and policy documents on competency-based assessment, we can also see that there is hardly any awareness of the needs of isolated learners. A review of recent conferences and forums (including the 1990 *Assessment and standards in vocational education and training* Conference, the two 1992 IIR conferences on competency-based assessment, the 1993 *National assessment research forum*, the 1993 *Testing Times* conference and the 1995 NTB *Global competencies—workplace outcomes* conference) showed that none of them had papers addressing these needs. Only passing reference can be found in a few journal articles to the potential that competency-based assessment (and training) might have to meet the needs of those living in isolated places.

Distance education has been an important part of education in Australia for some time. Primary education can be learned through the 'School of the Air', government and private schools offer boarding facilities for secondary education, and distance education and open learning have been actively pursued in tertiary education (Hall 1990, Lundberg 1993). However these have generally been simple extensions of the formal *face-to-face* education approach rather than *innovative* (or new) developments that are sensitive to the particular needs and limited resources of isolated learners.

Validity is the first requirement of good assessment. Multiple-choice questions from item banks are easy to administer to remote learners (Kleinschafer 1992) and standardised tests can be interpreted with powerful statistics (Johnson & McCaw 1992), but what do these tests really measure? While the tests do help with reliable assessment in remote locations and do allow for individual differences in pace, these approaches reduce flexibility and responsiveness to local needs and have doubtful validity (Docking 1991). Harris (1987, pp.78-79, p.132) warns of this danger of depending too much on reliability and not enough on validity in assessment in open learning.

The *reliability* of assessments for use in remote locations is the next important requirement. Concerns about reliability can include examiners assessing people differently and often leads to centralising the assessment process or putting in mobile assessment centres. There are other ways of improving reliability that avoid these problems. Rein (1993) describes Tourism Training Australia's strategy to 'present the competency standards in ways which would be consistently interpreted by trainers, assessors and candidates' to improve the reliability of the assessment process in their ACCESS system.

Flexibility is another requirement of competency-based assessment that is very important for remote candidates. Misko (1994, p.24) wants the kind of flexible assessment that will allow for variation in 'when, where, how and by whom, assessment is carried out'. This, along with flexible delivery, would 'open up the training system to students who, because of distance, work duties or family commitments, cannot attend courses presented on campus at times which are convenient' (p.15).

Finally, *fairness* in assessment is of great concern to remote learners, particularly those who are also culturally remote. Tania Ka'ai (1992) looks at the potential that *standards-based* assessment has for the Maori community:

> *Standards-based assessment recognises how important assessment is in defining the attitudes students take towards their work, their sense of ownership and control, the strategies they employ in their learning and their confidence and self-esteem—all of which impact profoundly on the quality of learning outcomes.*
>
> *Standards-based assessment complements the [Maori] philosophy—Students are assessed on what they know and not on what they do not know; students do not compete against each other, instead they strive to reach described levels of achievement thus encouraging co-operative learning; student performance is assessed in relation to specific criteria providing clearer goals thus identifying students strengths and the levels of achievement can be used to build up a comprehensive profile capturing the social context of learning.*
>
> (Ka'ai 1992)

We can easily apply this idea to Aboriginal communities, or any other groups or individuals for whom traditional approaches to assessment are inappropriate or inaccessible.

The ACCESS system for the recognition of competencies in the hospitality and tourism industry, developed in 1991 by Dean Ashenden and Russell Docking, was the first competency-based assessment system designed to reach isolated learners. The hospitality and tourism industry has a very dispersed workforce, in which many work in isolated communities in a wide range of establishments (Toop et al. 1994). Responding to this, the industry developed an approach to assessment that allows the widest access to workers in the industry, without losing rigour or quality. ACCESS not only helps employees in the workforce, but can also be used by those not currently working in the industry. The ACCESS system is in keeping with the design requirements listed on page 123 and places particular emphasis on establishing an extensive network of workplace assessors and industry assessors.

Even though ACCESS is 'the most developed skills recognition system in Australia' (Toop et al. 1994, p.60) the workforce has been slow to accept the system. It has been difficult to market assessment systems that are not directly linked to providing training.

Other problems, such as credibility and cost, are addressed in the following section.

Issues and challenges

The potential of competency-based assessment for remote learners

Competency-based assessment offers a unique opportunity to meet the needs of remote learners. In particular it allows learners to set their own pace, be assessed in their own environment and learn competencies in their own way—all without compromising standards. It concentrates on outputs—*what* has been achieved, however and wherever it was achieved; and not on inputs—*how* and *where* learning took place.

Competency-based assessment sees that a competency is the same, no matter how or where it was achieved. An element of competency is uniquely defined by its performance criteria and when these are observed together, that is enough evidence of the element's achievement. Similarly, a unit of competency is defined by its elements and when there is evidence of the achievement of the elements, some judgement can be made about achievement of the unit. Units can also be measured directly through *holistic* or *integrated* assessments that measure many elements at once (Gonczi 1994, p.29).

Integrated assessment offers real advantages in efficiency because assessments can cover many elements at one time. It also makes the assessment more relevant because the elements can be assessed together, showing 'the complex combinations of attributes' used to function in a particular situation (Gonczi 1994, p.29). Unfortunately because holistic or integrated assessments depend on professional judgement and experts directly observing learners (Gonczi 1994, p.29, pp.33-34) they will cost too much for remote learners.

Measuring whole units of competence has other problems. Units of competence (including, for example, tasks or training modules) are also likely to be very specific to their particular context and not of lasting relevance. Apart from those undertaking specific studies which deliver intact units of competency, remote learners will have little opportunity to practise or demonstrate similar units to learners in the city and will have little opportunity to use the units beyond their current context. If the assessment is to have more than local relevance it will need to be set up to recognise and record the achievement of *elements* of competency (Docking 1993, Rutherford 1995).

The basis for remote learners having their achievements recognised, therefore, needs to be *elements of competency*. These can be observed independently, transferred across different subjects, be widely applied and have a relatively long life. Records of the completion of tasks or training modules will therefore need to be interpreted in terms of the task or module completed as well as the elements that underpin them.

To make sure that these records are not just lists of isolated elements, they should include evidence from tasks or training modules of elements that show whether a person can link them with other elements (Docking 1995).

Competency-based assessment can be used in any context. It can be used with formal education and training to recognise achievements through distance education or external studies; it can be used for workplace training and recognition for remote employees (as with ACCESS); and it can draw upon life experiences and competencies learned during leisure time, as in the recognition of prior learning.

The problems of competency-based assessment in remote locations

The credibility of remote assessments

A major concern about assessment in any context is its credibility. This concern is very important when considering assessment in remote locations. What controls on quality can be provided? How can we be sure that the assessment is the same as assessments elsewhere? Can we be sure that no-one has cheated?

To some extent these concerns are addressed by the ability to separate evidence gathering from decision making in competency-based assessment. If evidence gathered at a remote location according to well-defined criteria can be authenticated by an independent and credible witness such as a teacher, clergy or police officer, then experts (or an expert system) can make use of that evidence to make a judgement and complete the assessment centrally.

The transferability of remotely-assessed competencies

People may also be worried about whether competencies can be transferred from one context to another context (Gonczi 1994, p.29). Concerns like these are increased if the competencies have been acquired in an isolated location, as many people might feel the learners have not had proper training. Transfer will depend on how clearly the competency has been specified (including the specification of acceptable evidence), particularly if nationally recognised specifications are used (or specifications that are accepted throughout Australia). However, it is always important to see that the achievement of the *component elements* of competency required for a task does not necessarily mean that the individual can co-ordinate them to actually *perform* the task. Therefore we may easily transfer elements, but find it more difficult to transfer units. This again reminds us that the performance of a task involves both the *elements* that can be measured anywhere and *integrating* (or bringing together) elements that might be unique to a particular task or context.

Isolated learners may acquire and demonstrate their competencies over a long period of time, or there may be a long wait between their assessment and recognition and subsequent use of their competencies. This raises concerns about the 'shelf-life' of competencies. Competencies (particularly units of competency) have limited shelf lives for a number of reasons. These include learners forgetting competencies, or changes occurring in the way they are used, particularly in areas of rapid technological change. This effect can be minimised by recording the achievement of *elements* of competency as well as *units* of competency, as elements generally have a longer shelf-life.

The costs of assessment

The cost of assessment in remote locations is a further deterrent. Costs include the specification of standards (although there is a growing resource of national standards with which to work), and establishing techniques to help gather evidence (these may include a network of assessors and/or validators and the development of assessment guidelines). Developing or purchasing software systems, office and communication services and monitoring systems can also be expensive.

These costs have to be balanced by benefits. For example, costs can be reduced by recording units rather than element achievement (Francis 1993), but this will prove costly in the end because it restricts the transferability and shelf-life of the competencies. Costs can also be reduced by avoiding a monitoring system, but this will reduce credibility and recognition outside the system.

Misko (1994) fears that the cost of adequately managing a remote assessment system 'might be more trouble than it is worth' (p.31) but we must be careful not to compromise validity in the quest for economy. We must avoid the temptation to use multiple-choice tests inappropriately, limit evidence, accept hearsay or second-hand evidence or the presumption of transfer (Francis 1993). Rather we should save money by using the learner to gather evidence and reducing the role of the expert to judging evidence. We can also formalise their decision-making rules through expert systems, or reduce travel costs by using technology such as facsimile or computer modem, or use a sampling process to monitor assessors and evidence gathering (Francis 1993). We could also employ phased assessment processes which use the cheapest forms first and only use the expensive forms *after* evidence shows they are needed (Bailey 1993).

The costs may be met by the user (the person being assessed), by employers, by industry groups or by training providers. Tourism Training Australia's ACCESS system charges a registration fee which covers transcripts of achievement and a handbook of standards. Extra fees are charged for assessments at assessment centres. Other fees are charged for assessors, which includes training, a manual and administrative support. Other costs such as assessor time are carried by their employer (Toop et al. 1994, p.57).

Cost-effective approaches

General requirements

Misko (1994, p.31) suggests that 'Involving candidates in their own assessment may free teachers, supervisors and managers from the heavy demands competency-based assessment makes on their daily work. It may also reduce the need to employ and train assessors and reduce the costs associated with having assessors regularly called away from their roles in production'. However, Misko is quick to remind us that this approach introduces potential new costs, such as the need for candidate training, the writing of specified standards, establishing and administering moderation procedures and the threat of mediocrity.

Some of these are transitional costs, but others will cost money over the longer term. In developing a cost-effective approach to competency-based assessment in remote locations, we will need to make sure that validity is preserved. This will require a careful matching of assessment method to the elements and units being assessed.

A cost-effective assessment method in widespread use is the short-answer or multiple-choice test. Such a test can be administered remotely by employing a local official such as a police officer, minister or public servant to administer the test under controlled conditions. Such a test, however, is only likely to show the learner's retention and recall of knowledge.

Other elements of competency that relate to thinking skills, psychomotor skills and attitudes will require more sophisticated evidence-gathering tools. The most appropriate strategies are likely to involve gathering a portfolio of evidence related to specific elements and signed by a local official (see Dugas & Horrocks 1993, Walton 1993). The evidence may be gathered from projects and assignments (Harris 1987, pp.88-89), leisure or work activities, or from formal studies. Other elements of competency may require attendance at assessment centres, the use of mobile testing units (Jensen 1985), or the use of new technology such as computers and satellite communication (see 'Use of technology' on page 151).

An example of a cost-effective assessment system for remote learners

The context
This assessment system was developed by a government department with a capital city central office and a large number of small outlying offices in remote country centres, staffed by one or two people.

Competency standards (units of competency, elements of competency and performance criteria) were defined by expert panels for a series of promotional levels within a single occupational group. Elements of competency were identified for over thirty distinct areas of expert technical knowledge and for units in the areas of: understanding the organisation, planning and organising work, working with others, finding and using information, expressing information, using technology and using mathematics and statistics.

Valid methods of assessment were found for each element of competency. Methods of assessment included written tests for knowledge, evidence written in a portfolio, interview and on-the-job observation. Existing training programs were also defined in terms of the competencies they develop.

The challenges
Some staff were intellectually isolated because they were not working in company with others who shared their particular expertise and many others were geographically isolated in small country centres. Existing staff needed to be assessed by the competency standards to see if they were eligible for promotion and to give information that would support the development of new training programs and training delivery.

Cost-effective and valid methods of using an assessment system had to be identified. As the area involved both safety and possible litigation, quality could not be compromised by *sampling competencies*. As the job was changing quickly at all levels, assessment had to be at element level rather than at unit level.

The assessment system had to be as accessible to those in remote locations as it was to those in the city.

A solution

A phased approach to assessment was used, moving from the most accessible and cheapest assessment methods to the least accessible and most expensive assessment methods. Each element of competency was assessed by the cheapest and most accessible valid method available. Each assessment phase used an integrated assessment which showed the learner's achievement of many elements of competency. An employee was not permitted to progress to a more expensive method of assessment until an acceptable level of performance had been achieved.

The phased system

Portfolio—The cheapest (for the organisation) and most accessible method of assessment was through the employee gathering written evidence of competency and preparing a portfolio. This meant that evidence of a wide range of elements could be brought together in any location. Costs to the organisation were limited to a half-day training program for all employees conducted during routine regional meetings, occasional advice and support from a training officer or line manager, and the review of the portfolio by an assessment panel. The employee had to link the evidence to one or more elements of competency and have it authenticated by an independent source. The portfolio was assessed by an assessment panel that included one person from outside the officer's region. The candidate could progress to the next phase if their portfolio showed enough evidence of achievement.

Written test—Multiple-choice and short-answer tests were used to measure technical knowledge. They allowed a sufficient sample of evidence to be quickly and cheaply gathered. Costs to the organisation included the writing, printing and administration of the tests. Sufficient evidence of achievement through the written test allowed the candidate to progress to the next phase.

Interview—The interview was used to gather two kinds of evidence. First, it was used to verify a sample of the evidence presented in the portfolio, and second, to gather new evidence that could be best found in that context. This included simulations of conflict resolution and flexibility. The main cost to the organisation was the time and travel costs for interview panel members. Interviews could be timed to coincide with regional visits or visits by the candidate to the central office. The gathering of sufficient evidence of achievement through the interview allowed the candidate to progress to the final phase.

On-the-job observation—By this stage the evidence required for most elements of competency had been gathered and the field observation could target those skills most useful for the workplace. This included communication skills, planning skills and the application (as distinct from the possession) of technical knowledge. Costs to the organisation included travel and time for expert observers.

Record-keeping

As evidence was gathered, a record of elements of competency demonstrated was collected. This record could then be scanned at any time to identify units of competency achieved and eligibility for promotion. Assessors could also be identified for panel membership or field observation for specific sets of competencies.

The records could also be scanned by region or by State to identify competency deficits and training needs. Records could build up information from this assessment process, or from other assessment opportunities including job interviews, training or performance appraisal.

Appeals

An executive committee with external assessment expertise managed the assessment system, monitored the work of assessment panels and heard appeals relating to assessment outcomes when required.

Future developments

New technologies, such as CD-ROM and networking are now being considered for the delivery of training to remote employees. These media could also be used to support the assessment process. Data-capture from normal computerised work logs could also be used to provide evidence for a number of elements of competence.

Future directions

Solutions and strategies for competency-based assessment for remote learners

The first problems that remote learners experience are lack of information about what services are available and how to get support from those services. Information tends to be distributed by industry, training provider or government agency. For example, to find out information about the competency requirements of a profession might involve contacting the National Office of Overseas Skills Recognition, the relevant professional body and a range of education and training providers. All this assumes that the applicant already understands which particular profession would be relevant to their experience.

A main requirement, therefore, will be a national *clearinghouse* holding all information, a 'one-stop-shop', which can counsel remote learners on assessment and training opportunities, support them through the assessment processes (and perhaps act as support person for the candidate) and then guide follow-up action.

Associated with such a clearinghouse would be the development of a national assessment *resource inventory* which would provide details of resources available at (or to) remote locations. For example, when they are in a particular location, the trained ACCESS assessors might be able to manage assessment outside the hospitality and tourism industry and current or retired teachers and lecturers might also help conduct assessments.

A range of communication devices may be used for gathering evidence:

- documentation (including mail, texts, portfolios and diaries);
- electronic communication (including telephone, facsimile, computer modem, video-conferencing and e-mail);
- electronic media (including cable television, computer disks, CD-ROM, smart cards and video);

- on-site assessments (by self, peers, supervisors, managers, independent invigilators, industry experts, mobile training units);
- off-site assessments (local industry, regional centres, assessment centres).

Finally the development of a national cross-industry record-keeping system—a *skill bank*—should be considered. This would enable remote learners to 'save up' evidence, elements and units for later recognition or use. In many cases remote learners move from centre to centre and the system would need to track such moves. The ACCESS system achieves this national coverage within the hospitality and tourism industry.

Competency-based assessment in remote locations

Implications for implementation

What are the implications of establishing assessment systems to meet the needs of remote learners? All the major players in competency-based training and assessment will need to consider the special needs of remote learners and modify their procedures to meet those needs.

It would be helpful to remote learners if employers described job requirements for recruitment, not only in terms of qualifications, duties and responsibilities, but also in terms of the elements of competency accepted as equivalent to those qualifications and necessary to be able to perform the duties and responsibilities. Remote learners would then be able to prepare more relevant job applications and compete more equitably with those with more traditional credentials. Employers may also consider on-site assessments of potential employees to assess their ability to apply their competencies, or offer employment initially on a probationary basis to allow for validation of the competencies and to observe their application.

Unions will need to accept that non-traditional skill development pathways can be equivalent to traineeships and apprenticeships and that competencies developed in one trade can be transferred to another. Likewise, professional associations will need to be able to accept that many of their profession's competencies are shared with others and can be acquired in ways other than professional education.

TAFE lecturers will need to describe their modules of study in terms of elements of competency, so that prospective students can be given credit for those modules or part of modules for which they can produce credible evidence of competency achievement. As with employers, TAFE lecturers are rightly concerned that the learner not only has achieved the component parts (elements) but can also tie these together to perform tasks or roles (units). Recognition that an applicant has evidence of all the elements may, therefore, be insufficient for the award of the module and a supplementary assessment might be required to assure the lecturer that the applicant can apply the elements in an articulated and integrated way.

For industry generally, the needs of remote learners impose further demands on the implementation of competency-based training and assessment. These requirements relate to the specification of standards, the decentralisation of assessment, the keeping and interpretation of records, and the provision of support and the dissemination of information to those in remote locations.

Conclusions

We have seen in this paper that there is very little research into assessment at remote locations and none at all on competency-based assessment in this context. While much has been written on training delivery through open learning or distance education, this usually fails to talk about assessment, or only looks for ways to replicate *on-campus methods*. There is a need, therefore, for a concerted research effort to explore the validity and effectiveness of the range of assessment options for remote learners and the infrastructure necessary to support them.

Such a research project would involve education and training providers, industry standards bodies, licensing authorities and national training authorities.

The needs of learners in remote locations will not be adequately met by the extension of formal training programs through distance education or open learning. Learners also need the competencies gained from everyday life to be recognised through credible competency-based assessment and for these recognised competencies to be linked to further training or work opportunities.

References

Atkinson, R, McBeath, C & Meacham, D, (eds) 1991, *Quality distance education ASPESA Forum 91*, Australian and South Pacific External Studies Association, Sydney.

Bailey, M 1993, *Judgement, evidence and the assessment of competency*, paper presented at the Testing Times conference, Sydney.

Docking, R A 1991, 'An A-Z of assessment myths and assessment in the workplace', *Competence and assessment* briefing series no.4, Employment Department Group: Standards Program, United Kingdom.

Docking, R A 1993, 'Hidden dangers in competency-based record keeping', in *Testing Times*, National Centre for Vocational Education Research, Adelaide pp.40-49.

Docking, R A 1995, 'Competency—what it means and when you know it has been achieved', *Global competencies—workplace outcomes*, National Training Board, Canberra.

Dugas, J & Horrocks, R 1993, 'Using the portfolio approach to assess Australian managers: A case study', in *Testing Times*, National Centre for Vocational Education Research, Adelaide pp.50-56.

Francis, J 1993, *Cost-effective competency-based assessment*, paper presented at the Testing Times conference, Sydney.

Gonczi, A 1994, 'Competency-based assessment in the professions in Australia', *Assessment in Education*, 1:1, pp.27-44.

Hager, P, Athanasou, J & Gonczi, A, 1994, *Assessment technical manual*, Australian Government Publishing Service, Canberra.

Hall, W C 1990, 'New directions in TAFE', in *Current issues in Australian education*, eds. J V D'Cruz & P E Langford, Longman, Melbourne.

Harris, D 1987, *Openness and closure in distance education*, Falmer Press, London.

Jensen, O 1985, 'Design and use of mobile units for technical and vocational education', *Studies in Technical and Vocational Education*, UNESCO, Paris.

Johnson, I & McCaw, B 1992, 'The South Pacific viewpoint on assessment and qualifications', in *Qualifications for the 21st century: International conference* 1992, New Zealand Qualifications Authority, Wellington pp.73-85.

Ka'ai, T 1992, 'A rationale for Maori assessment', in *Qualifications for the 21st century: International conference 1992*, New Zealand Qualifications Authority, Wellington pp.252-264 .

Kleinschafer, B 1992, 'Flexible Learning in WA', in *Australian Training Review*, no.4, pp.10-11.

Lundberg, D, 1994, 'Flexible delivery', in *Selected and edited papers from the National TAFE senior executives' conference 1993*, National Centre for Vocational Education Research, Adelaide.
Misko, J 1994, *Flexible delivery*, National Centre for Vocational Education Research, Adelaide.
Rein, A 1993, 'Workplace assessment: The ACCESS experience', paper presented at the National Assessment Research Forum, Sydney.
Rutherford, P 1995, 'Who's doing what? Assessment in Australia today' in *NTB Network*, no.16, National Training Board, Canberra, pp.19-20.
Toop, L, Gibb, J & Worsnop, P, 1994, *Assessment system design*, Australian Government Publishing Service, Canberra.
Walton, K 1993, 'A portfolio approach to development and assessment', in *Testing Times*, National Centre for Vocational Education Research, Adelaide, pp.185-190.

Author

Russell Docking is principal consultant of Skill Resource Management Systems Pty Ltd, a consulting firm specialising in competency identification and the development and application of competency-based assessment methods and systems, RPL, certification, competency information management and related human resource management systems. Dr Docking has worked as a teacher, as a teacher educator, as a senior public servant and as a consultant in all levels of education, vocational education and industrial training and has been instrumental in the introduction of many competency-based reforms in these areas. He has been a regular contributor to conferences, workshops and publications on competency-based assessment. He is currently involved in conducting national workshops on competency-based assessment methods, implementing computer-based competency-based record-keeping systems and providing consultancy services to a number of public sector and private sector enterprises throughout Australia.

Moderation

John Foyster

The idea of moderation in the context of competency-based training and assessment

The idea of moderation arises because assessment is a practice which is essentially imperfect—there can be no perfect and accurate assessment system. Those who believe that an assessment system (especially their own assessment system) can be or is perfect will always have difficulty with the concept of moderation and other devices that correct faults.

In the assessment literature the two words used in discussing faults which occur in the assessment process are *validity* and *reliability*.

In the present context an assessment fault due to problems with *validity* arises when the assessment system asks the 'wrong' question or series of questions; when the assessment system performs inconsistently we refer to problems of *reliability*. These different faults occur essentially because we invoke sets of rules—even agreed sets of rules—according to our own interpretations.

To remove or at least minimise these faults is critical in the competency-based context; there has been a lot of effort put into carefully designing competency statements and other components of the competency-based approach to training, and the value of that effort is much reduced if the assessment of the outcomes of training is faulty. Substantial effort should therefore be invested in making sure that the assessment component is as faultless as possible.

But improvement in the quality of an assessment system comes at a price. *Moderation* is a fault-correction device designed to achieve improvement at the lowest possible cost. This lower-cost improvement is achieved through a range of devices (examples will be described below), depending on the form of moderation employed, but the common device found in all forms of moderation is *sampling*. The role of sampling in moderation is summarised in *Designing a moderation system*:

> *Sampling enables effective use of resources that are available for moderation. Sampling involves targetting selected, representative points to check the quality of the whole.*
>
> (NZQA 1992, p.15)

In this paper we will look at the use of various forms of moderation in the context of competency-based assessment. The paper will initially deal with those aspects of assessment which lead to faults, and show which devices might be introduced to correct—or at least *moderate* those faults. We will then look at current and past moderation practices, both local and overseas, and discuss the suitability of these practices in the context of competency-based assessment. Finally cost effectiveness and possible future developments will be explored briefly.

A semi-technical introduction: How faults creep into assessment and what is done about them

Those setting up assessment systems understand that the systems can have faults, so reasonable steps are taken to minimise the size and effect of the faults. For example, we choose assessors who are capable of conducting assessments according to strict *criteria* or rules, possibly providing them with specific training. Careful steps are taken by assessment planners to establish standards which will be used by assessors.

These steps are widely understood and used. But it was recognised many years ago, particularly in the context of assessment at the university and secondary school level, that, even with such processes in place, faults still existed in the system which would produce seriously erroneous assessments.

The steps which were described in the first paragraph of this section take place *before* the assessment occurs. (In other words, they are designed to reduce the possibility of serious faults occurring in assessment.) But it is also prudent to include fault-correction steps in any assessment system. Moderation is one example of such a step, designed to occur *after* the assessment has taken place.

The move towards a competency-based approach in Australia has seen great investment in 'before the event' devices, which establish an environment in which good assessment practices can produce reliable and valid assessments. Whether sufficient effort has been given to the less glamorous side of assessment—cleaning up after problems have occurred—is much less clear. Moderation and other fault-correction devices which attempt to improve validity and reliability of assessments seem to be less well documented. As a result, in this paper we will describe the wide range of tested moderation devices. We need to look at so many approaches because the moderation device chosen has to be set in its context since, as we said earlier, moderation is an 'after the event' device intended to give significant improvement at a reduced cost.

Let's consider a contrasting fault-correction device: the appeals mechanism. Here the person whose work has been assessed and is dissatisfied with the result, can go through a formal procedure which allows some re-assessment of the situation by one or more third parties. Such a system depends on the aggrieved party coming forward, is extremely expensive, and is used only if the participant takes the first step.

A second procedure, also now less widely used, is the supplementary (or extra) examination or assessment. Here the initiator is the assessor who, within guidelines, decides that for certain participants there is (say) not enough evidence to decide finally on the outcome. In this situation the participant is offered the opportunity to undergo another assessment. This procedure is also expensive and operates on a case-by-case basis.

Other error-correction procedures try to anticipate the possibility of errors arising not just for some participants but for all participants. For example, an examination system may double mark papers submitted by participants (thus getting two opinions on every participant's performance), or may specify that different parts of an assessment be carried out by different markers. This process makes sure that the overall decision about a performance is made by more than one person.

These approaches to fault correction, which are useful in particular situations, are inherently different from moderation, which deals strictly with assessment outcomes. Moderation attempts to deal, after the event, with one possible source of error, and addresses that source of error directly.

In terms of the words used at the beginning of this section, moderation addresses directly the question of reliability (and only indirectly deals with validity) for those situations in which the lack of reliability *is due solely to the behaviour of the assessor. Put bluntly, if you believe that assessors occasionally make mistakes in their assessments then you will want to have a moderation system in place.*

Validity problems can be addressed by appropriate standards-setting exercises (which incidentally also improve reliability). In combination, moderation and the setting of standards attempt to deal with many aspects of the two problems of validity and reliability. (This is not to say that moderation cannot be used to improve the validity of assessments, which in some forms it can; but rather that other techniques will be more efficient.)

Moderation and verification

In this paper we consider that moderation is one component (or part) of the assessment system which may contribute towards the verification of assessments. (In other words, moderation will help confirm or check an assessment). Moderation is a form of verification with particular attributes, the most obvious of which is the emphasis on sampling.

Toop, Gibb & Worsnop (1994, pp.124-5) describe the role of verification in the assessment process, pointing out that verification processes may occur before, during, and after an assessment is made; while moderation is the main verification process which happens after the assessment is made. These authors also point out (p.28) the range of design options available in the development of a verification system, and their range of options and the reasons for choosing those options includes the possibilities for moderation systems. Moderation is, however, an all-or-nothing proposition—either there is a formal moderation system in place and it is universally used (within an assessment structure), or there is no moderation used at all.

In the rest of this paper we will occasionally refer to verification when it particularly applies to moderation.

Moderation—its meanings and limitations

We have identified moderation has been identified as a mechanism which addresses the problem of reliability, using sampling methods which lead to a reduction in costs. What does this mean in practice?

The faults which moderation seeks to correct are those introduced into the assessment system by the marker or assessor. It is assumed that assessors, for whatever reasons, make mistakes and the moderation system is designed to correct or allow for some of these mistakes. As a fault-correcting device, moderation systems are designed to address consistent faults on the part of the assessor, not random errors or certain kinds of particular errors.

- Moderation systems do not try to solve the problems which arise because *an assessor is biased towards or against particular candidates.* These are problems better handled by an appeals mechanism, or prevented by appropriate training.

- Nor do moderation systems claim to prevent every single error, however small, which may occur. Moderation systems, it will be recalled, are used as a low-cost device to reduce the number of faults, not to be a (necessarily) high-cost system which eliminates *all* faults.

These lower costs are deliberately achieved by monitoring only part of the assessment activities. The moderation system is designed to sample some carefully chosen components of the assessment system and check that faults found within those components are within agreed limits (which may be either fixed or loose). From such a check it may be inferred that the whole assessment system is behaving within the design specifications. Clearly such an inference involves some leap of faith. This partly explains why so many different moderation systems have been developed, as system designers attempt to produce schemes which will suit the particular belief systems of assessors in different walks of life.

Is the idea of moderation widely understood and accepted? In practice, the answer is no. Before pursuing the details of moderation systems this problem should be explored.

Lack of acceptance of moderation as a practice occurs mainly for two (possibly three) quite different reasons.

Moderation is unnecessary

First, moderation may be rejected as being unnecessary by those who believe that *the establishment of appropriate standards leads to completely reliable and objective assessments* by individual members of an assessment panel. This is a belief which can only be maintained for a short time—until the believer is confronted with concrete evidence of the unreliability of an assessment. 'True believers' will never, of course, acknowledge any unreliability, preferring to believe that their own assessments are foolproof.

Published evidence on the unreliability of assessments written for the lay reader is relatively uncommon, since such proceedings are normally treated as confidential. As a result of this, the belief that moderation is unnecessary is more widespread than might be expected by those who have studied the evidence. A recent experimental study (Wolf, Burgess, Stott & Veasey 1994) actually publishes two concrete examples which should convince all readers since both of the examples show considerable marker unreliability (although one example is a 'high agreement' group). A book-length study by Spencer (1979) of the problem in the school context is more detailed and may be useful to those wanting a technical discussion.

It is interesting to note in the case studies described by Toop, Gibb & Worsnop (1994) the high proportion of cases in which verification arrangements are seen as unnecessary (p. 53-113). Wolf, Burgess, Stott & Veasey (1994), in their study of the grading system for General National Vocational Qualifications (GNVQs), say that 'In many cases [verification] procedures have been developed only well into the first year of delivery. . . . A situation where no verification procedures are in place is obviously a transitional phenomenon, associated with a new award'. It may be a general phenomenon that verification (and thus moderation) is only seen as necessary in the light of experience,

and that accordingly the wide use of moderation will occur only after those managing assessment systems have had enough experience.

Moderation doesn't work

Second, particular moderation systems are rejected as unsatisfactory because while moderation may be recognised as necessary, the particular system in place does not give the outcomes the objector wants. As we saw earlier, this is why there are many moderation systems, as designers struggle with a range of desired outcomes.

Once again, there is little published evidence, but this is because objections are often overcome by the development of a 'new' moderation system. Later we will describe the outlines of several approaches to moderation; in fact, a comprehensive recent sketch of moderation methods has been published in New Zealand (NZQA 1992) which includes 12 examples, but is by no means complete at a technical level.

'There is a better (perfect) system of moderation which should be used instead of (this one)'

This is a variation on the previous case; it is also possible to identify a group of objectors who formally agree with the *principles* of moderation but in *practice* want very elaborate and expensive moderation mechanisms which go far beyond the modest aims of moderation systems. This group might be regarded as 'true believers' in moderation, since their theory tends to be that under their proposed moderation system' assessment does become infallible.

Since moderation systems can *always* be improved, it is tempting to change an in-place system to accommodate criticism. Improvements are almost always very expensive when made incrementally. Indefinitely prolonged 'improvement' can only be checked by firm financial controls which, in practice, assessment authorities have often found very difficult to impose.

It is important when designing a moderation scheme to steer between these attitudes, and accept modest outcomes from a moderation scheme: improved reliability at a low cost.

Moderation and panel assessment

There is one practice that we will *not* consider as moderation for the purposes of this paper. This is the use of a panel of assessors, rather than a single assessor, for each assessment. This practice is excluded from further consideration because it is not low cost (and the sampling, if any, is weak).

A system of panel assessment *is acceptable* in an assessment system when very accurate and reliable outcomes are needed and low cost is not an important factor. A widely-known example of one form of panel assessment is the use of panels of judges in such sports as diving and gymnastics, in which the final judgement about a competitor's performance is reached by eliminating the two most extreme grades (highest and

lowest) given by two judges and averaging (or totalling) the remaining grades. The result is then used to make fine distinctions between competitors (perhaps as accurate as one part in a thousand). Even so, controversy sometimes surrounds such judgements.

In the educational arena, such panel assessment arrangements (usually only involving two or three judges) are sometimes regarded as moderation systems. This is inadvisable for two reasons: the element of sampling which we identified as a key element of moderation) is very weak or does not exist, and the cost of such practices, relative to genuine moderation practices, is very high. An informal and unpublished study of the costs of Year 12 assessments in South Australia in the late 1980s showed that the cost of assessing a single participant in a subject could vary by a factor of *fifty* depending on the subject; *most of the high-cost variation occurred due to the use of panel assessment in the name of moderation.*

Accordingly, as a rule of thumb, it might be appropriate to limit the use of the term 'moderation' to systems whose total cost is less than double that of assessment by a single assessor.

Moderation systems in use

What exactly happens when assessments are moderated? Moderation systems can be classified according to the approach they use to sampling.

Sampling can be done in terms of:

- selecting a sample of each candidate's performance for further consideration;
- selecting all the work of a sample of candidates; or
- a combination of the above.

In practice, these distinctions are not frequently used (though they will be used below). More commonly, although many methods of moderation have been described (see, as suggested above, NZQA 1992), in fact moderation systems fall within one of three methods:

- statistical moderation
- visitation moderation
- group moderation.

These three forms of moderation are dealt with in turn below.

We will look at the general nature of each form of moderation and then explore the particular problems which arise in applying that form of moderation to competency-based assessment. After the three forms have been outlined in this way, we will further consider their possible use in competency-based assessment.

1 Statistical moderation

Statistical moderation is a method in which (most commonly) a sample of each participant's performance is used. In practice, how the sampling is done is usually determined by the kind of assessment which occurs and the way in which the original assessment is conducted.

Consider, for example, the most widespread form of moderation used in the Australian secondary school system—widely and almost generically known as 'statistical moderation', although it is actually only one example of the method. (The technical method of carrying out this form of moderation varies slightly from use to use, but the principles are essentially the same.) It should be noted that this example is given not because it is appropriate but because it is the system most widely used in Australia.

Any system of statistical moderation assumes that each assessment involves working with a sample of each participant's performance, and that there are at least two such samples. In the Australian Year 12 systems, these two samples are typically an external examination and school assessments by the teachers of the participants. In the context of competency-based assessment, such an approach could be used where assessment of both skills and underpinning (or supporting) knowledge was regarded as essential; and the assessment of underpinning knowledge was through a common (national) form of assessment while the skills assessment was conducted locally.

The two assessment components exist because it is believed that either component alone is not a large enough sample of the participants' performances. In combination, it is believed, the two components produce a better summary of each student's performance. Statistical moderation is formally needed because the total assessment system, as described above, is not sufficiently controlled.

In this form of moderation it is assumed that there *is* sufficient quality control of the common assessment through the participation of many people in the production of each individual assessment, but insufficient control of the local assessments. Accordingly, in this model of moderation, the common assessment can be and is used as a moderating device.

What must be moderated is the grade given by the assessor. The reason for moderation is not that the local grade is inherently defective or wrong, but rather there has been no opportunity for adequate cross checking. Control of the local assessments is achieved statistically (hence the name of this form of moderation) by placing statistical limitations upon the grades which may be given by each assessor submitting local assessments, based upon the grades awarded in the common assessment. (Variations in what these statistical limitations are produces the range of methods referred to above.) These statistical limitations are not applied to compare each participant's common assessment grade with her/his local grade, but only to the whole set of grades submitted by each assessor; a participant-by-participant comparison would constitute a re-marking of the local assessments.

Although there are variations between systems and subjects, in practice the overall result of this system of statistical moderation in the schools context is that the school-assessed grades are slightly lower after moderation and also are spread out slightly more. There are technical reasons for this pattern which we will not describe further here.

A key characteristic of statistical moderation which partly explains its popularity is that its incremental cost over the initial costs of carrying out the assessments is almost nothing; statistical moderation is carried out at very low cost by relatively simple computer programs which can operate on data already entered into data banks.

General applicability in the context of competency-based assessment

Statistical moderation as we have specifically described it above has limited applicability to a system of assessment which is competency-based.

1 Like all statistical methods, this form of moderation requires a relatively large number of participants to be meaningful. For example, if there are fewer than ten participants in the group whose assessments are to be moderated then applying most forms of statistical moderation cannot be defended.

2 Many forms of statistical moderation rely upon an assessment scale which includes many grades (e.g. a 100-point scale). Most forms of competency-based assessment depend on a much shorter scale (e.g. a two-point pass-fail scale).

3 Finally, as indicated above, the use of this approach requires the existence of a form of common assessment across all participants (see below for further details).

If statistical moderation is to be used in an assessment system which is competency-based then it must be a form of statistical moderation which can operate with a short assessment scale and for small numbers of participants. In general, this eliminates the forms of statistical moderation most commonly used for Year 12 assessments throughout Australia.

However, we should not be assume that assessment scales in competency-based systems are so short that statistical moderation is completely ruled out; Hager, Athanasou and Gonczi (1994, pp.85-93) forcefully describe the case for the role of grading scales of modest length, summarising their view as 'grading is possible and, in certain circumstances, desirable in a competency-based system' (p.93).

What forms of statistical moderation are available which might be relevant? First, we can name the forms of assessment which might be used as the moderating device. The NZQA publication identifies *external examinations, common assessment tasks, reference tests, and item banks*; each of these assessment forms is produced by a panel of assessors independently of the on-the-spot assessor whose assessment is to be moderated.

Development of each of these forms of assessment is relatively expensive and must usually therefore have already occurred before this form of moderation could be considered. External examinations and common assessment tasks, for example, may stand alone as components of an overall assessment system, as may (though this is less common) reference tests and item banks.

The key difference between the first and second pairs of systems is that reference tests and item banks have been tested thoroughly and have quite accurately-known characteristics (for example, it may already be known that an item bank is answered correctly by 65 per cent of the relevant population), whereas information about *how difficult external examinations and common assessment tasks are* is known only after they have been conducted.

All four forms tend to result in assessment extending over a long range (though an item bank consisting of pre-tested performance tasks may shorten the scale). Item banks, reference tests and common assessment tasks may be constructed to consist of easily identifiable components which address particular competencies and could therefore theoretically be adopted for a competency-based system. However, this is not the way such forms of assessment are usually constructed.

Furthermore, this moderating device must be matched with the assessment by the local assessor which will almost certainly be on a short scale. The problems in routinely using the four methods we have described above as statistical moderating devices, are therefore considerable. Is there anything here which deserves serious consideration?

Statistical moderation is generally regarded as the cheapest form of moderation and, despite the considerations above, it does deserve further investigation. Some form of statistical moderation will be useful in maintaining standards in a competency-based scheme and will offer a low incremental cost provided that:

- the moderating device has already been developed and is in use;
- the problems of the small size of assessment groups can be solved;
- the need for a short assessment scale can be accommodated.

2 Visitation moderation

Usually a sample of each participant's performance is used in this form of moderation, although it is possible to use the whole performance of a sample of participants if there are enough of them. In this case the moderating device is not a common assessment, but an external moderator, who visits the assessment site (hence the name). In a true visitation moderation procedure there is a statistical element, as will be shown below.

Three forms of visitation moderation can be identified.

In the purest form of visitation moderation an expert moderator visits each assessment site and considers the assessments given by the local assessor. The expert moderator also considers a sample of performances by the participants, and then *uses statistical guidelines* to decide whether the assessments given by the local assessor fall within acceptable limits. If they do not, either an appeals mechanism may be used or the expert moderator may give advice to the central authority that a re-assessment is required.

In addition, where an appropriate atmosphere exists, local assessor and expert moderator may negotiate the group of assessments. *If they negotiate all the assessments then this occurs as a second assessment, not as a form of moderation.* Suppose the assessment system requires assessments of fail, pass, and credit. The expert moderator may argue that the available evidence suggests there should be three credits, five passes and two fails, rather than two credits, five passes, and three fails as the local assessor believes (in other words, the expert thinks the local assessor is too harsh); they may then agree to split the difference—two credits, six passes, and two fails, using generally-agreed guidelines.

(It should be noted that such a form of moderation does not require the local assessor to give the same grades as the expert moderator, merely that the differences be reasonably small.)

A second form of visitation moderation occurs when the situation described above applies except that the 'visit' is *virtual* as a result of the use of mail or electronic media. (In other words, the moderator does not visit in person). Since a major part of the cost of visitation moderation is travel, virtual visitation becomes increasingly attractive, although there are limitations on the kinds of performance which can be moderated in this way.

A third form of visitation moderation is much weaker than the two examples we talked about previously. In this form the expert moderator visits the local site merely to discuss assessments, standards, and general procedures, perhaps with bench mark examples where this is possible.

General applicability in the context of competency-based assessment

Such a system of moderation, although expensive, has attractive features for a competency-based assessment system.

1 It can serve as an informal mechanism for maintaining standards amongst assessors and for bringing them up to date with recent developments in the field.

2 It is easily applied for both short grading scales and specific competencies.

But this form of moderation tends to be expensive, since the expert moderators have to be hired either full time (for large-volume activities) or released from other senior responsibilities (for small-volume activities), on top of any travel costs. Where assessment sites are far apart the only realistic option may be virtual visitation moderation, which may not be satisfactory for some performance tasks.

Also it may not be easy to obtain the services of appropriately expert moderators, upon whom the system absolutely depends.

Consideration of visitation moderation therefore implies a substantial financial commitment for the single purpose of moderation (bearing in mind the possibility that there may be spin-off in the development of local assessor expertise). Nevertheless, there may be situations in which this form of moderation can be justified.

3 Group moderation

Where statistical moderation and visitation moderation use only external means to achieve the objective of maintaining consistency, group moderation weakens this external influence somewhat. In this form of moderation the checking is carried out by peer assessors, rather than an external expert moderator.

Once again, three variants on the basic form can be identified.
In the pure form, sometimes known as *consensus moderation*, a panel of peer assessors meets and, using an agreed set of procedures, considers various performance samples from assessments made by individuals within the group. It must be emphasised that if this process is to be regarded as moderation then only a sample of performances can be considered; assessment (or re-assessment) of *all* the work of all the participants is a *full panel assessment*.

The judgement to be made, here as elsewhere, is whether the range of assessments made by each local assessor is within the acceptable range. Two special requirements of such a system are (a) a very carefully designed set of procedural rules (or standards of behaviour) which are widely used and (b) appropriate rules defining the behaviour of the local assessor when her/his assessments are being moderated.

The two following variants are different enhancements of consensus moderation. The reason for having two variants is a desire for more widespread consistency, since if too many panels exist there can be variation in practice between the panels.

In the first variant the panel of assessors is enlarged to include an external expert moderator. This addition can in some circumstances lead to a moderation system which is in fact visitation moderation by an expert in the presence of a panel, but this need not be so. This variant is capable of great range, from a form in which the expert moderator actually makes the decisions to a form in which the expert moderator goes no further than helping the panel.

In the second variant the local panel includes at least one person who is also a member of another panel. In this way a network is created which, if appropriately managed, will achieve reasonable consistency across the whole network. We should remember at this point that moderation is intended to achieve reasonable improvement for reasonable cost, and thus a compromise approach of this kind may be most attractive.

General applicability in the context of competency-based assessment

This system of group moderation also has attractive features for a competency-based assessment system.

1 Like visitation moderation it can serve as an informal mechanism for maintaining standards amongst assessors and (in some circumstances) for bringing them up to date with recent developments in the field.

2 It fosters the development of formal and informal networks which may have general benefits.

3 It is also easily applied in terms of both short grading scales and specific competencies.

There are, however, some disadvantages in using this approach. Although it is inherently less expensive than visitation moderation, it is not so rigorous, and may be regarded as inadequate in some situations. The need to bring the panel of assessors together may be difficult to achieve when members have to be released from work, and impossible when the assessors live far away from each other.

Using of this form of moderation, although attractive in some ways, might therefore be limited to particular circumstances.

What moderation cannot do

Before describing how the three forms of moderation might best be used in verifying (or confirming) judgements made in a competency-based assessment system, two particular limitations should be made clear.

First, as talked about earlier, moderation is not designed to produce a 'perfect' assessment result. The very idea of moderation comes from the belief that assessments by one person are not satisfactorily reliable. If an individual's judgement is somewhat unreliable, then it does not follow that the decision of a group of individuals will be (by contrast) completely accurate. It only means that, in the right circumstances, the group decision will be more reliable than the individual's decision. Those who want a perfectly reliable assessment system should not look to moderation systems as an answer.

Second, there are two forms of assessment for which moderation systems really cannot be developed.

Since moderation as we have defined it here is dependent upon taking a sample of each participant's work in the set of participants assessed by each assessor, there are considerable sampling difficulties in the case of *self-assessment* since there is only one participant and one assessor who happen to be the same person. In such a case the general rules surrounding the conduct of the assessment have to be relied upon for verification purposes.

The same sampling problem occurs in the more general case where the one participant and the one assessor happen to be different people. Thus *peer assessment* as a form of assessment presents difficulties for anyone designing a moderation system (and the same problem exists for any situation in which an assessor has only one participant). It may be possible to design a moderation system to cope with these situations, but the difficulty and cost may be too great to make it worthwhile.

What moderation can do

The three forms of moderation we have talked about in this paper can all play a useful part in improving the reliability of competency-based assessments. What kind of moderation is most appropriate depends upon the circumstances under which the assessments are conducted.

The table on the next page summarises some key system features in any consideration of the use of moderation procedures as described above.

General moderation type	Specific moderation example	Cost issues	Implementation issues	Special requirements
Statistical moderation	Using external examination	Moderation is *low cost per capita* if external examination already exists.	Requires large number of participants; suitable for dispersed participants.	Common across all participants. Centralised assessment. Extended grading scale.
	Using common assessment tasks	Moderation is *low cost per capita* if common assessment tasks already exist.	as above	Common across all participants. Extended grading scale.
	Using item banks	Moderation is *low cost per capita* if item banks previously developed and standardised.	as above	Common pattern of use by all participants. Extended grading scale.
	Using reference tests	Moderation is *low cost per capita* if reference tests are already developed.	as above	Common across all participants. Extended grading scale.
Visitation moderation	Expert visiting moderator	Moderation is *high cost per capita.*	Effective for small numbers of assessors, provided they are not too dispersed.	Requires appropriate experts.
	Expert 'moderator at a distance' (virtual moderation)	Moderation is *medium cost per capita.*	Effective for small numbers of assessors.	Requires appropriate experts. Requires assessment to allow consideration at a distance.

General moderation type	Specific moderation example	Cost issues	Implementation issues	Special requirements
	Moderation by expert (discussion)	Moderation is *high cost per capita.*	Effective for small numbers of assessors, provided they are not too dispersed.	Requires appropriate experts.
Group moderation	Consensus moderation	Moderation is *medium cost per capita.*	Effective if significant numbers of assessors are in close-proximity clusters.	Requires experienced assessors, more preliminary training.
	Networked panel moderation	Moderation is *medium cost per capita.*	Effective if significant numbers of assessors are in close-proximity clusters and linking assessors can network.	As above
	Panel moderation with expert	Moderation is *high cost per capita.*	Effective if significant numbers of assessors are in close-proximity clusters.	Requires experienced assessors, available experts.

In summary the significant features of the three main moderation types may be described as follows.

Statistical moderation is the least expensive form of moderation. It may be used in limited circumstances because it:

- normally requires at least ten participants per assessor (if fewer than ten participants, strong statistical techniques will be required);
- normally requires a reasonable number of points on the grading scale (five or more)
- can only be used where an inexpensive moderating device already exists (such as an item bank, common achievement task or reference test or other common assessment).

Because statistical moderation is inexpensive, it should be the form of moderation of choice where the requirements are met. Choice of specific method depends on which moderating devices are already available and upon the number of participants per assessor.

Visitation moderation is the most expensive form of moderation in per participant terms.

There are circumstances in which visitation moderation may nevertheless be the best form of moderation, as follows:

- sufficient expert moderators must be available;
- assessment sites must be physically accessible to external moderators.

If visitation moderation is to occur 'virtually', the performances being moderated must be of a form which can be examined at a distance.
Although visitation moderation is expensive, it may be the most practical approach in some situations. In particular, it will be complementary to other forms of moderation.

Group moderation generally falls between the other two forms in terms of cost per participant. It has particular advantages in certain circumstances as follows:

- where the assessment sites are clustered;
- where the development of links between assessors is regarded as a substantial side benefit;
- where a new system is being introduced and professional development is regarded as a priority.

In summary, there are no hard and fast rules which suggest one universal system of moderation; the most appropriate form of moderation is determined by the circumstances under which the assessments are being conducted. But the range of existing and tested forms of moderation is adequate for most assessment circumstances.

Moderation—what needs to be done

The benefits of using moderation systems have been outlined above. Several key next steps are implied.

First, although the need to moderate assessments by local assessors is so well established in many areas of educational assessment that it is accepted as essential and understood without question, this is not yet universally the case in competency-based assessment in the vocational education and training area. Relevant research in these other areas of education is now over 20 years old (and therefore to some extent out of date), while research in the VET area (for example, Wolf, Burgess, Stott and Veasey [1994] as cited above) is only beginning the task. While the work of the kind reported by Wolf, Burgess, Stott and Veasey (1994) could profitably be continued in Australia, older work in other areas of education should also be consulted in order to identify new and relevant directions for research methodologies.

Second, the special assessment problems which make moderation (and other components of an assessment system) difficult in Australia should be investigated in detail. The major problems are the relative diversity of approaches to assessment and the high dispersal of the population over a large physical area; these problems have high cost implications. It is probable that extended use of new approaches involving sophisticated information technology may alleviate some of these problems and reduce costs of assessment systems.

Finally, since moderation is intended to add value to assessment for relatively low cost, but nevertheless implies added cost, users of assessment systems must be convinced of the merit of this 'added value'. Research conducted into assessor variability and moderation methods must accordingly be sound, thorough, and well-marketed.

Conclusion

It appears from both experiments and case studies that people do not immediately recognise the importance of moderation when assessment systems are being devised. However, as time passes, and people gain more direct experience of the outcomes of the assessment system, the need for appropriate verification is more widely accepted. Moderation, as a form of verification designed to provide significant improvement at low cost, is likely to be identified as important, as economic considerations start to have an effect on the design of second and later generation assessment systems.

The move towards using moderation systems will produce better assessment sooner. However, because there is no 'universal solution' time is needed for system designers to identify the best path forward.

Just which moderation method should be used depends upon a range of factors as described above. There are combinations of factors for which satisfactory moderation methods at reasonable cost have not been adequately developed, and the most productive area for research in moderation will be the consideration of further methods and their applicability in the wide range of circumstances to be found in which competency-based training occurs in Australia. This paper has outlined options for rational improvement in the arena of competency-based assessment.

Bibliography

Hager, Athanasou, Gonczi, 1994, *Assessment: Technical manual*, AGPS, Canberra.

NZQA 1992, *Designing a moderation system*, Wellington.

Spencer, 1979 *Folio assessment or external examination*, Scottish Certificate of Education Examination Board, Edinburgh.

Toop, L, Gibb, J & Worsnop, P 1994, *Assessment system design*, AGPS, Canberra.

Wolf, A, Burgess, R, Stott, H & Veasey, J 1994, Institute of Education, University of London, London.

Author

John Foyster is the manager of the statistics division of the National Centre of Vocational Education Research Ltd. He has a BSc (Hons), Monash and a broad background in research, evaluation and statistical procedures. Prior to joining NCVER, John was co-ordinator (Evaluation and Curriculum Development) at the Senior Secondary Assessment Board of South Australia, a teacher and senior school co-ordinator at Preshil, Victoria and a senior researcher at the Victorian Institute of Secondary Education and the Australian Council for Educational Research. His recent publications include 'Selection for vocational education: Some lessons from elsewhere', in W Hall (ed.), *What future for technical and vocational education and training*, NCVER, Adelaide 1993 and *Getting to grips with on-the-job competencies*, TAFE National Centre for Research and Development/DEET, Adelaide 1991.

Use of technology

David Foreman and Russell Docking

Scope

Technical terms are in the glossary at the end of the paper.

This paper provides the reader with an overview of how computers are being used to support competency-based assessment, particularly in Australia. Implementation issues are looked at and cost-effective approaches are described. As this area is rapidly developing, some predictions about the future uses of computers in assessment are made.

There are few assessment situations which can be totally based on computers. However, in many assessments, computers are used for part of the process. Examples are provided to illustrate how and where computers are being used.

The term 'competency-based assessment' refers to a process of gathering evidence and making judgements from that evidence relative to appropriate competency standards, or course/program learning outcomes in the absence of standards. A 'computer-based competency-based' assessment is one where all or some of the stages are supported by the use of computers and related technology.

Thomson (1993) describes an approach which is both 'high-tech' and 'high touch':

> *computer technology applied to data gathering and record keeping has the potential to release trainers and administrative staff to give more time to individual trainees on a personal level.*

Without inexpensive computers, competency-based assessment would have remained a good idea that was never fully realised. In particular, computers are needed to store detailed information about competency achievement and for analysing those records for differing purposes.

The stages of competency-based assessment include:

- specifying competency standards (including units of competency, elements of competency and performance criteria) or learning outcomes;
- specifying assessment strategies (including item banks);
- selecting and producing assessment strategies;
- identifying and preparing assessors;
- gathering evidence (during learning, including adaptive testing; after learning, including conventional tests; through work activity, simulation or recognition of prior learning (RPL));

- scoring or scaling evidence;
- recording evidence;
- interpreting competency achievement (to make decisions relating to further training, employment or certification); and
- monitoring and reviewing the assessment process.

Not all of these stages are always involved in the competency-based assessment process, but computers can link each stage together. In the next section, examples of software are described to illustrate each use. By referring to selected software products, most of which have installations in Australia, the authors do not imply that it is the only software available, nor do they imply any endorsement of the software.

Australian and overseas practices

The need to use computing technology in competency-based assessment has been recognised in Australia and overseas for some time. In referring to competency-based and self-paced learning in TAFE colleges, Watson (1994) states:

> *In the experience of this author, some system of computer managed learning and assessment appears to be the only efficient and cost effective method of providing adequate records of progress whether a college employs self-paced learning or not. (p.36)*

This need for records management has been the primary focus for computer applications in assessment, and there are many examples of this type of application.

Australian examples of the use of computers in assessment are centred around record-keeping and the development of 'item banks' of relevant information. There are also instances in which computer-based assessment is included in computer-managed learning and with equipment simulators.

Others have proposed a far wider role for computers in assessment. A UK survey explored possible requirements for computer-based knowledge and skills testing, simulation and fault-finding, RPL, assessment decisions and record-keeping. While record-keeping was the most common use, there was significant interest in the use of computer-based materials and tests (Ward 1991).

Computers have been used in integrated training and assessment in several Australian organisations for some years, for example, RAAF purchased a maintenance training package with its purchase of FA-18 planes. It included computer-based assessment at the end of each session. This package was developed prior to today's PCs, but is still being used and updated.

For each of the nine components of competency-based assessment described above, there are examples of computer systems that support and use them. In some cases the systems go well beyond managing the assessment process, to training planning and management and/or human resource management. In the following discussion we will focus on assessment applications. Further details of the example software applications are provided in the box following, and a table, relating software to the assessment steps is given at the end of this section.

Examples of software systems currently used in Australia

System:	ASTEC
Publisher/distributor:	
Operating system:	286 PC
System:	CAAS (computer assisted assessment system)
Publisher/distributor:	B Kerwood, ACM Training Centre
Operating system:	286 PC
System:	CBTS-CML
Publisher/distributor:	CBTS (Australia)
Operating system:	VAX
System:	DATA PERFECT (ACCESS)
Publisher/distributor:	Tourism Training Australia
Operating system:	
System:	ELMSOFT
Publisher/distributor:	WA TAFE External Studies
Operating system:	286 PC
System:	INTRAIN
Publisher/distributor:	ICI Botany
Operating system:	IBM PC
System:	MicroCAT
Publisher/distributor:	
Operating system:	286 PC
System:	ScanMark
Publisher/distributor:	Scanform Computer Systems
Operating system:	DOS Windows 386 PC
System:	SKILLMAP/KIDMAP
Publisher/distributor:	Skill Resource Management Systems Pty Ltd/Mercator Software Pty Ltd
Operating system:	DOS 386 PC/DOS Windows 386 PC
System:	TestMaker
Publisher/distributor:	Scanform Computer Systems
Operating system:	DOS Windows 386 PC
System:	XAUTHOR
Publisher/distributor:	DAT Corporation and The National Building & Construction ITC Ltd
Operating system:	386 PC

Specifying competency standards or learning outcomes

Computers can be used in specifying competency standards in the following ways:

- by providing ways of developing competency standards;
- by managing vast amounts of linked information (such as units, their elements and their performance criteria);
- by providing ready access to details of requirements; and
- by making it easy to create and edit links between competency standards and specific occupational or training requirements (such as job descriptions or qualifications).

The Tourism Training software (see page 153) based on DATA PERFECT for its ACCESS system presents competency standards in an easy to follow format to ensure that assessors and assessees fully understand the performance requirements. XAUTHOR provides a means of developing and editing competency standards. SKILLMAP/KIDMAP manages the units of competency, their elements and performance criteria, eliminating bulky paper work and providing infinite flexibility. Decision-making rules also can be incorporated, linking element achievement to unit achievement and unit achievement to job or qualification achievement. CBTS-CML can store course materials, objectives and structure.

Specifying assessment strategies

Computers can be used to specify competency-based assessment strategies by:

- providing standard text editing screens for standard assessment formats (such as multiple choice tests, project guidelines or observation checklists);
- incorporating graphics, illustrations or animation;
- providing instructions to assessees or assessors;
- storing assessment strategies in an item bank; and
- linking assessment strategies to learning sequences or workplace processes.

ELMSOFT, TestMaker, CAAS, ASTEC, MicroCAT and CBTS-CML each allow the system manager to enter and catalogue questions either directly or through a word processor. Questions can be catalogued by taxonomies for later selection. CAAS and TestMaker have the ability to include graphics in the questions. MicroCAT allows the developer to insert text, graphics and digitised images for simple conventional tests or for complex adaptive testing. ASK allows assessees to 'test out' their skills and knowledge before formal assessment.

Selecting and producing assessment strategies

Computers also can assist competency-based assessment by:

- allowing assessors to select assessment strategies to match competencies and workplace or learning contexts;
- printing assessment materials or checklists;
- displaying questions or tasks on-screen for the assessee;
- selecting further questions for the assessee based on answers given.

Four of the selected systems offer this feature. These are ELMSOFT, TestMaker, ASTEC and CBTS-CML. Tests can be developed by randomly selecting items using preset taxonomies or item-by-item selection. Tests are then printed or are displayed on the screen (CBTS-CML). MicroCAT allows tests to be printed or adaptive assessments to be administered on-line. Bennett (1994) suggests that in computer-adaptive implementations, tests can be considerably shorter than their paper-and-pencil counterparts, without sacrifice in measurement precision'. ASK provides questions at three levels relevant to office skills competencies.

A program has been developed for the national occupational health and safety certification system allowing the random selection of items from specified groups. These include observation checklists, assignments and short answer assessments. Knowledge-based items can be produced in alternative formats to suit oral and written assessments. Graphics have not been included to allow for the use of older computers.

Identifying and preparing assessors

Competency-based assessment requires careful matching of assessor with assessee. Computers can assist by:

- maintaining a record of assessors and the competencies they are accredited to assess;
- matching assessors with assessees by competency, location and time;
- monitoring the effectiveness of assessors;
- supporting the development of assessors.

Tourism Training's application of DATA PERFECT retains information about assessor competencies, areas of special strength and localities so that workplace competency-based assessments can be arranged for candidates.

Each OHS authority in the national OHS certification system is establishing computer-based registers of assessors and the classes of equipment they are registered to assess. These registers will be used to provide assessees with lists of assessors, as well as for the auditing of assessments.

Gathering evidence

Computers can be used to gather evidence of competence. This includes:

- presenting standard multiple-choice tests to be answered using the keyboard or pointing device (mouse);
- presenting adaptive assessments where the computer is programmed to select a sequence of questions dependent on the answers given to each;
- supporting the work of assessors by recording their observations;
- monitoring work output linked to competency achievement; or
- simulating work situations to monitor assessee response, particularly where dangerous, expensive or rare activities or products are involved.

Most of the selected systems (ELMSOFT, ScanMark, CBTS-CML, CAAS, DATA PERFECT, INTRAIN, MicoCAT, SKILLMAP/KIDMAP) provide a means of accumulating assessment evidence, but only CBTS-CML and MicroCAT can do so by

directly interacting with the assessee. This software also has the potential to be linked to computer aided instruction, and CAAS can be interfaced directly with a wide range of digital tools and equipment.

Several simulators have been purchased recently by Australian organisations to train and assess crane and mobile plant operators. It has been reported that at least one cost in excess of one million dollars. There are attempts to develop hardware and software solutions within the price range acceptable for Australian industry, with a target price range of $100 000 for a simplified plant operator system (O'Brien 1995). This type of simulator is being used in the building and construction industry around Australia. Less expensive simulations in the form of interactive games may also be useful as assessment tools.

Computer recording of work output has been used for training and assessing people in computer-based work such as data entry and transaction entry, direct sales and stores work. This type of application requires the full co-operation of the assessees and other employees if it is to be acceptable.

Scoring or scaling evidence

Computers can help with processing evidence in order to:

- determine when sufficient evidence has been accumulated for competence to be recognised; and to
- scale responses against standards where appropriate.

ELMSOFT, CAAS, MicroCAT and ScanMark can be used to score tests and to assign grades.

Recording evidence

The computer is often used for recording the results of assessment. Electronic marks books which convert assessment information to marks and tally these have long been available. Competency-based assessment requires that very detailed records of the achievement of elements and units of competency are maintained. Course accreditation and trainer registration requires that learner records are retained for a number of years. Detailed records of elements and units can also be used to support later decision-making and RPL.

Most computer-based competency-based systems provide for the storage of performance information, including ELMSOFT, ScanMark, CBTS-CML, CAAS, DATA PERFECT, INTRAIN, SKILLMAP/KIDMAP. ScanMark offers scanned data entry to feed data into an electronic marks book. ELMSOFT, SKILLMAP and DATA PERFECT offer structures for storage that match the Australian National Training Authority (ANTA) competency structures. ASK records answers for analysis against clerical and administrative competency standards.

Interpreting competency achievement

Computer-based records of elements and units of competency can be interpreted to:

- identify group training needs;
- evaluate training programs;
- plan work allocation;
- plan further individual training;
- award certificates or licences;
- recognise prior learning; or
- assist career planning.

Standards would need to be set by experts, including patterns of competencies required for the performance of each function.

All systems provide some form of analysis of individual performance ranging from listing competencies achieved (DATA PERFECT and INTRAIN), to grades (ScanMark) to complex analysis of performance and standards matching (CAAS and SKILLMAP/KIDMAP). Some systems provide the facility to store comments from assessors for other assessors or for assessees (SKILLMAP/KIDMAP and CBTS-CML). ASK provides the assessee with a measure of progress against the national clerical and administrative competency standards.

Group performance can also be analysed, from simple score distributions (ScanMark and INTRAIN) to complex performance analysis and profile based ranking (CAAS, CBTS-CML and SKILLMAP/KIDMAP).

Monitoring and reviewing the assessment process

Computer-based records of competency achievement can be used to identify inconsistencies and other indicators that will assist in:

- monitoring assessor performance;
- improving the clarity of competency standards;
- verifying the relevance of expert systems; and
- validating the usefulness of assessment methods.

The data stored in ELMSOFT, CBTS-CML, CAAS, ScanMark and SKILLMAP/KIDMAP can be interpreted to monitor the effectiveness of the assessment process. ScanMark provides standard item analysis options. The Assessment Systems Corporation offers a range of item analysis programs for use with its test software (MicroCAT and ASTEC).

The preceding sections illustrate the extent to which competency-based assessment is being supported by computer systems. Some of these merely use the computer to replicate paper and pencil tasks, while others take advantage of the speed, capacity and flexibility of computers to extend competency-based assessment into new areas.

Kerwood (1993) and Ward (1991) report that computer-based competency-based assessment can contribute to the efficiency, effectiveness and quality of the assessment process. There have been a number of reported adverse comments about the use of

computers in assessment. The authors have found from experience and observation that computer-based assessment is well accepted when there:

- is full consultation with assessors and assessees on its use;
- are clearly identified advantages;
- is avoidance of the 'big brother' image by providing adequate personal support and access;
- is inclusion of assessees in the computer-based assessment process.

Other concerns and issues are examined in the following section.

Software system applications

		AS	CA	CB	DA	EL	IN	Mi	SC	SK	Te	XA
1	Specification of competency standards or learning outcomes			✓	✓					✓		✓
2	Specification of assessment strategies	✓	✓	✓		✓		✓			✓	
3	Selection and production of assessment strategies	✓		✓		✓					✓	
4	Identification and preparation of assessors				✓							
5	Gathering of evidence		✓	✓	✓	✓	✓	✓	✓	✓		
6	Scoring or scaling of evidence		✓			✓		✓	✓			
7	Recording of evidence		✓	✓	✓	✓	✓		✓	✓		
8	Interpretation of competency achievement		✓	✓	✓		✓		✓	✓		
9	Monitoring and review of the assessment process	✓	✓	✓		✓		✓	✓	✓		
		AS	CA	CB	DA	EL	IN	Mi	SC	SK	Te	XA

AS: ASTEC
CA: CAAS
CB: CBTS-CML
DA: DATA PERFECT
EL: ELMSOFT
IN: INTRAIN
Mi: MicroCAT
SC: ScanMark
SK: SKILLMAP/KIDMAP
Te: TestMaker
XA: XAUTHOR

Implementation issues

Using computers in competency-based assessment offers great potential to promote competency-based systems. There are, however, a number of major issues that must be looked at, including:

- developing suitable technology;
- recognising the limits in the range of skills that can be computer assessed;
- links with training;
- links to human resource management;
- access to the systems, particularly in remote locations;
- industrial issues;
- privacy and security of information;
- user computer skills and comfort;
- costs.

These are discussed in the following sections. Ward (1991) provides a comprehensive study of the application of computers to assessment.

Developing suitable technology

Using computer-based competency-based assessment effectively totally depends on the availability of suitable technology. This includes hardware and software.

There are two options: to purchase software developed by others for particular hardware, or to develop a system for a specific use. The latter approach may cost up to $500 000 and three years development time. In addition, special skills in assessment and computer technology need to be combined if the system is to be 'user-friendly' and applied successfully.

Existing packages may cost between $2 000 and $20 000, and should be assessed against the system requirements. Plans should be made to overcome our problems which may affect successful implementation. For example, a system may not be able to hold all of the required information such as complex graphics or video, yet the cost to incorporate additional system features may be too high. In this case a manual approach needs to be developed in conjunction with the computing systems.

The application needs to have either of the following if it is to succeed:

- stable competency formats and standards;
- flexibility to incorporate updates to standards.

It is also desirable to have standard and inter-connectable data formats and computer platforms, with appropriate use of communications technology.

Any chosen system should have a flexible database, to allow for different types of assessment instruments, and other information. The most flexible forms of software are those with a modular design.

Recognising the limits in the range of skills that can be computer assessed

It is important to recognise that computer-based assessment is limited in its ability to measure skills. Clearly it can initiate formal tests by presenting text and graphics to the

assessee. Computers can also process answers, especially multiple-choice responses, but also simple text answers. Computers are also able to simulate activities (such as driving a car or flying an aircraft) but while this can provide a means of assessing competence, it should not be used in isolation. Care must be taken to match the assessment method to the skill being assessed and to avoid diminishing a skill by the assessment method. Ward (1991) reminds us that 'computer assisted assessment . . . should not be regarded as a substitute for observation of the candidate in workplace activities, and it cannot be expected to provide a complete solution to all the problems identified' (p.77).

Links with training

Assessment is an integral part of training, and computer-based assessment systems should be designed to complement and enhance the teaching and learning process. Computer-based competency records may provide a better guide to training needs and student selection. Continuous computer-based assessment can be used to provide the basis for continuous curriculum design, and computer-based assessment may offer improved training evaluation.

Links to human resource management

Recording the achievement of competencies can be used to assist many human resource management functions, including work redesign, skills audits, recruitment and selection, succession planning, performance management, career planning and redeployment. In turn, a number of human resource tools, such as selection procedures, performance appraisal and in-house training can provide evidence of the achievement of competence. Although not necessary, a computer-based competency-based assessment system that can link with the human resource management system or produce reports that support human resource decisions should be considered.

Access to the systems, particularly in remote locations

Learners in remote locations are likely to find competency-based assessment less available. Using computers and related technology (such as e-mail) for competency-based assessment could do much to overcome isolation (see also *Competency-based assessment in remote locations* in this series of papers).

Industrial issues

The use of computers raises a number of industrial concerns. These are dealt with in the following sections, but it is important to recognise that where computer-based competency-based assessment is being used there must be agreement on how the data are to be used, who can use the data, and how data quality is monitored.

Computing technology may add to the fears already held about assessment. The systems need to be clear about the standards required, the assessment methods to be used, and the records decisions obtained from the assessments.

Privacy and security of information

Of major concern is privacy and security. On the one hand the systems have to be sufficiently accessible and user friendly, but at the same time they must preserve personal privacy and prevent the misuse of data. Computer-based competency-based software systems will need to use passwords, user traces and perhaps encryption (refer glossary) to maintain security. Hardware security and the use of stand-alone (rather than networked) systems can also improve security.

Managers of assessment data need to ensure the information is only used for purposes with which the assessees have agreed at the time of assessment.

User computer skills and comfort

The software must not prevent the full use of competency-based assessment systems. Large, clear screens, Help systems, effective function-driven manuals, hot-lines and on-screen training should be supplied with the software.

Costs

Expenses associated with using computers in competency-based assessment include the purchase costs of the software and hardware, installation costs, costs of populating databases, and running costs. These costs have to be measured against the benefits.

Both direct cost-savings and indirect benefits should be evaluated to determine the benefits of computer-based competency-based assessment (see the next section for further discussion).

The issues outlined in this section may seem to present obstacles to introducing computer-based competency-based systems. Each of these concerns can be resolved, however, and there are cost-effective approaches available. These are discussed in the section that follows.

Cost effective approaches

There are a number of reasons why computing technology is being used in competency-based assessment. The justifications for its use in the examples reviewed include:

- cost savings;
- improved quality of records;
- improved quality of assessment;
- consistent data collection and analysis, (and therefore reliability of outcomes);
- safety (e.g. assessing competence in emergency situations);
- support for self-paced learning; and
- random item selection.

The rapidly reducing cost of computing hardware has increased the potential uses of computer-based competency assessment. However, the cost of software developing has not been reducing at the same rate, and more sophisticated assessment applications are relatively expensive.

The specific costs of using computer-based competency-based assessment are:

- system review and selection;
- hardware purchase and operation;
- software purchase/development and ongoing maintenance; and
- resources required to load the assessment materials into files/databases.

Experience has shown that the last item is frequently overlooked, and the materials needed for assessments in most courses are extensive. In one such case, the software was developed and ready for use, but the cost of loading the assessors' documentation onto the system was so expensive that the project was halted.

Where computers have been appropriately used there is general agreement that the validity and reliability of the assessments have improved.

The suggested approach for any organisation wishing to use computer-based competency-based assessment is as follows.

- Contact your State/Territory industry training advisory board to obtain:
  - details of the industry assessment procedures: it is desirable that your procedures are compatible with the industry procedures for information and credit transfer
  - obtain details of the existing and proposed industry use of computers in assessment: in particular determine if your organisation can obtain right of access to the industry systems.

- Review the industry procedures and systems against your requirements to determine if they can be used or modified to suit your use.

- If it is determined that the organisation needs a computing system, ensure that the requirements are carefully specified and that all costs of purchase, installation and operation have been included.

- Carefully review the alternative systems before purchase as their features vary substantially. No one system has a complete range of features.

- After installation be prepared to explain fully the system to assessors and assessees to overcome any resistance and fears.

- Allow sufficient resources so that the assessment system remains up-to-date, as the validity and reliability of assessments can be eroded by out of date information.

Throughout, take care to involve all parties in the decision-making processes.

In assessment, where additional costs are incurred through the use of manual or computer-based techniques, the issue of who pays generally is raised. At present, most assessment costs related to off-the-job training are incorporated into the training provider's charges. The main motivating factors for providers to use competency-based

learning and assessment systems have been:

- to reduce costs, especially in record-keeping;
- to improve the quality of training and assessment, especially in regard to consistency;
- to provide more self-paced learning and assessment.

With the introduction of competency-based training and the establishment of industry training advisory boards, practical performance in the workplace is required in many situations to establish competence. Careful planning is advised, so that costs are kept to a minimum.

It is probable that the costs of these types of assessment will be borne by the workplace. In addition, some workplaces have established recording systems, and costs have been passed on to trainees or employers. A computer-based assessment which can reduce time and other costs, therefore will be desirable.

For example, in computer-based work activities, computer logging of activities completed can be used as assessment evidence. This has been done successfully in a stores situation, and has been proposed for areas such as business planning and sales. The costs of assessment are restricted to collation of the evidence while useful work is completed.

Future directions

Current use of computers in competency-based assessment is limited. One survey of Australian TAFE colleges showed only a small amount of assessment records was being kept on computer, and this appeared to be the most common application (Friendenfelds 1993). Ward (1991) in a survey in the United Kingdom found extensive interest in computer-based assessment, but a limited number of uses. Ward states:

> *It is notable that almost no use is made of computer delivered assessment in assessment for national vocational qualifications, despite predictions of the potential for using computers in assessment which have been made over the past 10 years. In part this may be due to the rapid changes both in computer technology and in qualifications which have discouraged development. It may, however, indicate that external funding is required if significant progress is to be made.* (p.77)

Four levels of computerised testing and assessment have been proposed by Linn (1993):

- **computerised testing**: administering conventional tests by computer;
- **computerised adaptive testing**: tailoring the difficulty or content of the next item or test based on the assessee's previous responses;
- **continuous measurement**: using measures included in a curriculum (or work) to continuously estimate the person's achievement and learning profile; and
- **intelligent measurement**: using knowledge banks and expert systems to score the assessee's progress, and to provide advice to the assessee and trainers or assessors.

From the examples given (see page 152), most applications are at the first two levels cited above. Considerable scope exists for computing and communications technology in the areas of case studies, simulation, continuous measurement, adaptive testing and

of expert systems and knowledge inference. Also there is potential to expand the capability and use of existing systems.

As discussed above, the development of computing technology has made the introduction of competency-based training and assessment more feasible. Other issues are discussed below. The first five deal with advances in hardware, the remainder with software developments.

CD-ROM and related technologies

This type of equipment is suitable for large amounts of data stored as text, graphics, animation, video and sound. It also cannot be easily corrupted, providing for consistency and for repeated use. The potential for this type of equipment may include:

- combined delivery of training and assessment;
- development of large item banks, including graphics, video and audio information;
- adaptive testing with many alternative branches;
- case studies, with realistic workplace information; and
- cost effective simulation.

Of this type of development, only WORM (write once—read many) type technology is appropriate for the maintaining of records.

Scanning, OCR and writing interpretation

These technologies can be used in three areas - storing assessment materials, scoring assessments and analysing assessment items. Typically, hand written or typed information is read and sensed electronically and processed. Processes include optical mark recognition (OMR) or optical character recognition (OCR).

The most common examples of OMR are the TAB and cross-lotto forms. Mark sensing of answers in various forms of academic or psychological testing has been used for many years (see also the SCANFORM system). It is highly suited to multiple choice questions, allowing rapid scoring and analysis of test results. It has been used for mathematics and science testing of primary students across Australia.

The OCR technology puts typed or written information into computer coded text. This can be stored, analysed and processed. OCR has the potential to sort assessment items, read assessees' answers and process them prior to scoring and analysis.

Newton writing pads have become available from Apple in the past few years. These use a variation of character recognition software to process handwritten information.

In the future, the increasing availability of writing interpretation software may allow for the analysis of free response answers to both short answer questions and essay type questions.

Scanning technology provides an alternative to using a keyboard. It allows one computer to process the results of many assessments, or to enter text rapidly and cheaply.

Voice recognition and synthesis

Voice recognition and synthesis technologies provide a way of overcoming difficulties in reading and writing in some situations, and also reduce the time required for oral assessment.

These technologies could be used to enhance user friendliness of the system.

High capacity systems

The availability of higher speed and capacity systems at moderate cost has made graphical presentations, delivery of audio and video segments, adaptive testing, simulation, animation and on-line monitoring of performance feasible. Each of these options can be used by software developers to provide more flexible and capable packages.

In turn, assessments can be produced that may be used in a range of ways, which can relate more closely to performance at work, and which are valid and reliable.

The availability of powerful portable computers at moderate cost allows for computer-based assessment in remote locations, such as in rural areas (see also *Competency-based assessment in remote locations* in this series). Communications technology can also be included to improve access to other data and to report results.

Communication systems

Most of the applications looked at are based on centralised or stand-alone computers performing a limited range of functions. Communication between these systems is based on paperwork or transferring information between computers.

Communications options, such as networks, bulletin boards and the Internet should be used to provide such services as:

- speedy transfer of information, such as updates to assessment items, communication between assessors and assessees and assessment results;
- access to sophisticated assessment systems;
- help for assessors;
- auditing of assessors.

This type of development could be introduced using various open learning systems technology currently being developed.

Multimedia

Software development in this area has been based on several of the hardware technologies mentioned above, i.e. CD-ROM, higher capacity systems and the integration of audio and video technologies. Use of this type of software will need to store the assessment data in a user friendly way.

Authoring software

A number of developers are marketing increasingly sophisticated packages for developing assessment items, structuring assessments and recording results. Effort is required to make these packages easy to learn and use, and to allow individual assessment approaches to be used.

Simulation software

On page 156, one case study is described of recent developments in computer-based simulators. More sophisticated software is being developed to simulate the characteristics of equipment and controls digitally and to utilise the increased computing power now available at moderate cost.

The cost of simulators will reduce when applications are found which will encourage larger sales of specially designed equipment associated with the simulators.

The technologies discussed above will not be used unless they meet a need in a cost effective manner. The commercial packages described in pages 2–8 have been chosen for a specific, identified purpose, often at some commercial or technical risk.

The more expensive and risky developments may need to be trialled before being used in a commercial package. Some of the developments suggested above may cost hundreds of thousands of dollars (ScanMark and TestMaker took three years and $500 000 to develop). Seed research projects will reduce the investment risks substantially.

All of the discussion above is based on a supposition that assessors and assessees will become more computer literate over time. Assessors, especially, must feel familiar and comfortable with the computing system to use it effectively.

It must be remembered that computer-based competency-based assessment will only ever be part of the assessment process. Bennett (1994) foreshadows a future where there will be five different assessment options, in order of increasing cost and complexity.

- Computer delivered and automatically scored. This would include multiple-choice responses and also other forms of response that could be fully digitised for analysis. These systems could be fully adaptive to modify the sequence of questions to match responses.
- Computer-delivered and semi-automatically scored. In this case, the computer would recognise responses it could not score and pass these on to a human judge.
- Computer-delivered and human-scored on computer terminal. This corresponds to the use of the computer for record-keeping and analysis.

- Conventionally delivered and human-scored on terminal.
- Conventionally delivered and conventionally scored by human judges.

Priority research projects are suggested in the following areas:

- profiles of assessors and their ability to use computers;
- the extent to which different forms of computer-based assessment might be viewed negatively;
- the extent to which computer-based assessment might reduce flexibility or distort assessment methods;
- limitations of current assessment authoring software and techniques for improvement;
- costs and benefits of using communications technologies in assessment;
- use of multi media-capable computers in case studies and projects and in scoring responses: this type of development has the potential to provide alternative methods of assessing performance;
- use of voice recognition and synthesis in assessment.

Conclusions

This paper has provided a brief look at a rapidly growing and increasingly complex area of technology. Most of the information on applications has come from industry and from the commercial providers of the technology. The reader is encouraged to explore these and other sources of information to develop an awareness of current uses and possibilities.

The potential for the use of computer-based competency-based assessment is widespread. However the current use of computers in most assessment situations is small.

The section on Australian and overseas practices demonstrates that computer technology can be a useful tool in all phases of the competency-based assessment process. At the same time, computers cannot completely take over the role of assessing students.

Assessment is the process of gathering and evaluating information and evidence and often requires an approach that only humans can provide.

A wide range of issues must be addressed if the use of computers in competency-based assessment is to be successful, as discussed in the section on implementation. Pages 10–11 describe some cost-effective approaches to implementing computer technology in competency-based assessment, covering the types of costs to be considered and the possible justifications of expenditure.

The potential for the future use of computers is explored in 'Future directions'. The challenges, and the opportunities, are huge. Bennett (1994) maintains that the technology is available for sophisticated assessment delivery, scoring and analysis (albeit in a fragmented and, in some cases, experimental way), and that the networks are available to disseminate such a system, but that these developments are unlikely to be realised without a 'multi-organisational effort'.

We therefore have proposed some areas in which research is required to realise this potential, and invite researchers and industry to participate in and further develop the effective use of computer-based competency-based assessment.

Bibliography

Bennett, R E 1994, *The role of technology in creating assessment that increases participation in post-compulsory education*, paper presented at the 20th Annual IAEA Conference, New Zealand.

Friendenfelds, U 1993, *Competency assessments and recording processes: 1993 TAFE survey*, Marleston College of TAFE, Marleston.

Kerwood, B 1993, 'CAAS: A practical assessment tool', in *Testing Times*, National Centre for Vocational Education Research, Adelaide, pp.142-150.

Linn, 1993, *Educational measurement*, third edition, American Council on Education, Onyx Press.

O'Brien 1995, *Improved training delivery by use of high fidelity simulators*, University of New South Wales.

Thomson, P 1993, *Report on overseas investigations*, paper presented at the National Assessment Research Forum, Sydney.

Ward, C 1991, *Potential for computer assisted assessment in the assessment of national vocational qualifications*, Final Report, Research and Development Series no. 6, Employment Department, Sheffield.

Watson, A 1994, 'Three strategies for the assessment of competence,' in *Australian TAFE Teacher*, vol.28, no 1, pp.35-39.

Glossary

'Adaptive assessment' is an assessment where the assessment sequence is modified according to previous answers given or demonstration of skills.

'Black box' is term used to describe systems and/or parts of systems where the operations are hidden from the users.

'CD-ROM' is short for compact disk—read only memory. It is a form of optical data storage on a disk like a musical CD.

'Compatible platforms' are where two computer systems can freely exchange data.

'Competency-based assessment' refers to a multi-staged process of gathering evidence and making judgements from that evidence against appropriate competency standards or course/program learning outcomes.

'Computer-based competency-based assessment' is where one or more assessment stages are supported by the use of computers and related technology.

'E-mail' is short for electronic mail and refers to systems that record, transmit, store and control messages, e.g. notes, memos and letters, between system users.

'Encryption' refers to the use of programs and codes to transform data and messages into non-standard formats to restrict access.

'Hard-coded' is a term used to describe placing of data into computer program coding.

'Note-books' are small portable computers, typically with IBM-PC or Apple compatible formats.

'Profile based ranking' refers to the use of numerical techniques to develop profiles of classes of assessees and assessment items.

'Records-management' is a term for storing and controlling records, often on a computer data-base.

'RPL or recognition of prior learning' refers to the acknowledgement of skills and knowledge obtained through formal training, work experience and/or life experience.

'Score distribution' is the plotting of assessment results by frequency for various scores.

'Simulation' refers to an assessment situation away from the workplace, where workplace conditions are 'simulated' for the assessee to demonstrate aspects of competency. Some workplace conditions can be simulated on computers.

'Taxonomies' refers to the grouping or classification of assessment items, from which alternatives are selected.

Authors

David Foreman completed an engineering degree before working in the manufacturing industry for over 20 years, in engineering and management positions. During this time, he completed a post-graduate diploma in business management. An interest in the development of people grew with the successful implementation of technology and organisational change. This led to employment as a consultant for almost three years, before commencing his own consultancy, David J Foreman & Associates. As a consultant, David has completed a number of projects at industry and enterprise level in the development of competency standards, competency-based training and assessment.

Russell Docking is principal consultant of Skill Resource Management Systems Pty Ltd, a consulting firm specialising in competency identification and the development and application of competency-based assessment methods and systems, RPL, certification, competency information management and related human resource management systems. Dr Docking has worked as a teacher, as a teacher educator, as a senior public servant and as a consultant in all levels of education, vocational education and industrial training and has been instrumental in the introduction of many competency-based reforms in these areas. He has been a regular contributor to conferences, workshops and publications on competency-based assessment. He is currently involved in conducting national workshops on competency-based assessment methods, implementing computer-based competency-based record-keeping systems and providing consultancy services to a number of public sector and private sector enterprises throughout Australia.