



**A study into the assessment and reporting of
employability skills of senior secondary students**

REPORT

To the Commonwealth Department of Education, Employment
and Workplace Relations

From the Australian Council for Educational Research
31 January 2008

Prepared by Gabrielle Matters and David Curtis

The period of this review was:
July 2007 to January 2008

2008 Australian Council for Educational Research (ACER)
ABN 19 004 398 145
ACER Brisbane Office
Suite 9, 541 Boundary Street, Spring Hill, Brisbane
PO Box 658, Spring Hill, QLD 4004, Australia

Contents

Executive Summary	i
Chapter 1 – Introduction	1
<i>Quality of Education Review Committee</i>	2
<i>Finn Review</i>	2
<i>Mayer Committee</i>	3
<i>Employability skills for the future</i>	5
<i>Australian Certificate of Education</i>	6
Related initiatives.....	6
<i>National Goals for Schooling</i>	6
<i>Enterprise and Vocational Education</i>	7
<i>Australian Blueprint for Career Development</i>	7
Emergence of employability skills in vocational education and higher education	8
<i>The VET sector</i>	8
<i>The higher education sector</i>	9
<i>Target groups of students</i>	10
Conclusion	10
Chapter 2 – Fundamental Concepts and Issues	12
Criteria for evaluating approaches to assessment and reporting	12
<i>Validity and related constructs</i>	12
<i>Reliability and related constructs</i>	13
<i>Objectivity</i>	14
<i>Feasibility</i>	15
<i>Usability</i>	16
Roles and purposes of assessment	16
<i>Purposes of assessment</i>	16
<i>The signalling function of assessment</i>	17
<i>Implications for assessment and reporting of employability skills</i>	17
Notions of generic, work-related and employability skills	17
Conclusion	21
Chapter 3 – Performance Levels	22
Note about completeness	22
Employability skills for assessment.....	23
Towards a set of described performance levels	27
<i>Illustration</i>	29
<i>Rationale for describing facet standards</i>	29
Performance levels and evidence.....	30
Two-stage process.....	30
<i>Analytical method for combining results</i>	32
<i>Holistic method for combining results</i>	32
Audit of processes used for setting, describing and representing performance levels.....	32
<i>Establishing levels is conditional</i>	34

<i>Form of representation chosen</i>	34
<i>Note on status of descriptors</i>	34
Evaluation of descriptors	34
Description of performance levels for an employability skill.....	35
Descriptions of facet standards for an employability skill.....	38
Representation of profile of results	40
Reporting	41
Conclusion	42
Chapter 4 – Models and Methods for Assessing and Reporting Employability Skills Achievement	43
Descriptions of assessment methods.....	43
<i>Standardised tests</i>	43
<i>Common assessment tasks</i>	43
<i>Teacher-generated tasks</i>	44
<i>Judgment by groups of teachers</i>	44
<i>Embedded development and assessment</i>	44
<i>Portfolio construction</i>	44
<i>Self-assessment</i>	45
<i>An employability skills assessment model</i>	45
<i>Fitness for purpose</i>	45
<i>Control of assessment</i>	46
<i>Degree of standardisation</i>	46
<i>Specificity of assessment</i>	47
A framework for classifying assessment methods.....	47
Possible approaches to assessment	48
<i>Assessment approaches and employability skills</i>	49
Application of criteria for evaluation assessment approaches.....	50
Reporting issues.....	53
Conclusion	54
Chapter 5 – Preferred Methods for Assessing and Reporting Employability Skills	56
Main strengths and weaknesses of assessment options.....	56
<i>Standardised testing</i>	56
<i>Common assessment tasks</i>	56
<i>School (or teacher)-selected (or -devised) assessment tasks</i>	57
<i>Teacher-group judgment</i>	58
<i>Embedded development and assessment</i>	58
<i>Portfolio construction</i>	59
<i>Self-assessment</i>	60
Consequences of assessing and reporting employability skills achievement	60
<i>Authority of reporting or certification</i>	60
<i>Comparability</i>	61
<i>Consultation</i>	61
<i>Professional learning</i>	61

<i>The development of related skills</i>	62
Viable approaches to assessing and reporting employability skills achievement.....	62
Stakeholder perceptions	62
<i>Common assessment tasks</i>	63
<i>Teacher-group judgment</i>	64
Possible implementation strategies	64
<i>Standardised testing</i>	64
<i>Common assessment tasks</i>	65
<i>Teacher-group judgment</i>	66
Conclusion	66
References	67
Appendices	74

List of tables

Table 1:	Summary of generic/employability skills developments in Australia, by sector	11
Table 2:	Employability skills – foundational and applied facets	24
Table 3:	What was involved in the production of a set of described performance levels	27
Table 4:	Options for reporting performance levels as grades	30
Table 5:	Application of design rules to the production of described levels of performance	33
Table 6:	Summary descriptions of possible assessment approaches.....	49
Table 7:	Evaluation of assessment approaches	51

List of figures

Figure 1:	Opportunities for the development, assessment and reporting of employability skills for students on various paths from compulsory schooling into the labour market.....	10
Figure 2:	Diagrammatic representation of overlapping sets of skills.....	19
Figure 3:	The identification of performance levels – three per skill, referenced to advanced, creditable and basic achievement.....	28
Figure 4:	Two-stage process for arriving at overall grade for a particular skill.....	31
Figure 5:	Diagrammatic non-hierarchical representation of employability skills.....	40
Figure 6:	Diagrammatic representation of results for reporting	41
Figure 7:	Characteristics of assessment.....	45
Figure 8:	Framework for assessment models	48
Figure 9:	Clusters of employability skills along cognitive–affective and knowledge level dimensions.....	50

Acknowledgments:

Jennifer Bryce, David Kelly, Robert Lake, Doug McCurry, John Pitman

ACER Internal Reference Group:

John Ainley, Margaret Forster, Geoff Masters, Phil McKenzie

DEST's Advisory Group:

(Including proxies who attending meetings)

Australian Chamber of Commerce and Industry: Mary Hicks

Australian Council of State Schools Org: Jenny Branch, Steve Carter

Australian Industry Group: Peter Canavan

Australian Parents Council: Caz Bosch, Ian Dalton

Australian Secondary Principals Association: Jim Davies, Graeme Oliver

Business Council of Australia: Patrick Coleman

The Association of Heads of Independent Schools of Australia: Allan Shaw, Barbara Stone

EXECUTIVE SUMMARY

Background

The structure of Australia's economy continues to evolve in response to increasing competition within and from abroad and in response to developments in information and communication technologies. Accompanying the structural changes in the economy is a need for a more highly skilled and flexible workforce. The influence of these factors on Australia's education and training systems has been recognised since the mid-1980s. The 2002 report by the Australian Chamber of Commerce and Industry (ACCI) and the Business Council of Australia (BCA), *Employability Skills for the Future*, again drew attention to the need for young people to acquire a set of higher-order cognitive skills—employability skills—before they leave school.

Subsequently, Masters, Forster, Matters, & Tognolini (2006: vii) recommended that 'schools and authorities give priority to developing and recognising young people's progress in relation to the eight employability skills' as identified and defined by ACCI and BCA. This recommendation is based on the assumption that raising the profile of employability skills would deliver higher levels of their achievement among new entrants to the workforce.

Objectives and supporting activities

In order to investigate the most effective ways of assessing and reporting on the employability skills of senior secondary students, the Commonwealth Department of Education, Science and Training (DEST) commissioned this report.

The main activity was to evaluate options for assessing and reporting on eight employability skills against five criteria—validity, reliability, objectivity, feasibility, and usability—and to recommend a preferred approach. This work was undertaken by the Australian Council for Educational Research (ACER) in the period July 2007 to January 2008, during which time the (new) Department of Education, Employment and Workplace Relations (DEEWR) became responsible for the initiative.

The starting point for this work was to come to terms with the eight employability *skills* and their respective *facets* from the *Employability Skills Framework* (ACCI & BCA, 2002). The employability skills are: Communication; Initiative & Enterprise; Learning; Planning & Organising; Problem Solving; Self-management; Teamwork; and Technology. Facets are elements of the skill that employers have identified as important, with the specific mix and priority of facets being job-dependent. One of the facets of Communication, for example, is 'Reading independently'.

As an adjunct activity to this study, the University of Western Sydney conducted a survey of the current level of employer satisfaction with the eight identified employability skills and how employers assess them (Costley, Power, Watson, Steele, & Sproats, 2007).

Consultations were undertaken with employers, and representatives of parent organisations, school systems including teachers and leaders, and Australia's three peak business organisations, ACCI, BCA, and the Australian Industry Group (AIG). An advisory group set up by DEST had the same representation. In addition, expert input was sought on detailed matters of assessment and reporting.

Outcomes and conclusions

Overview

Assessing and reporting on employability skills achievement in a way that would conform to the hallmarks of educational measurement (e.g. validity and reliability) is a highly complex task, which places a high—even unrealistic—level of demand on the education system. And so there is the possibility that rigorous assessment requirements could lead to abandonment of the goal as not worth the effort. It would be a major achievement, however, merely to instil employability skills into student learning so that transition from compulsory schooling to the workforce is more efficient and effective from any perspective. Reporting achievement is important and, while a Rolls Royce may be the desirable vehicle, it may be unobtainable. Some progress, however, is at least possible with a Citroen 2CV. Less than rigorous assessment may be the necessary price to pay for a general improvement in the employability skills of Australia's young people. This point about compromise should not be lost.

Through an analysis of examples of implementation of generic and employability skills within Australia and from overseas, a set of major approaches to the development, assessment and reporting of employability skills was elicited. The criteria for evaluating these examples were themselves scrutinised, as were the possible purposes served by different assessment approaches.

Criteria for evaluation

Traditional definitions of the five evaluative criteria—validity, reliability, objectivity, feasibility and usability—were found wanting when they were applied to those assessment approaches identified as possibilities for employability skills. For example, objectivity is not simply a property of the final stage of assessment (awarding a grade to a student). Judgments are made at all stages of the assessment process from selecting the form of assessment, choosing tasks and contexts, writing prompts, establishing standards and criteria, and grading work. Further, the criteria have differential influences, depending upon the main purpose to be served by the assessment.

Purposes of assessment

Of particular interest in senior secondary schooling is the enhancement of the employability skills of individuals and of the cohorts of young people who follow various pathways in transit from compulsory schooling into the workforce. This purpose, served by some assessment approaches, is in addition to assessment that leads to fair, reliable and objective reporting of achievement. These purposes—encouraging development and facilitating reporting—need not be exclusive; indeed, good assessment can support both purposes. Stobart (2004) states that 'effective assessment encompasses the dual and related goals of formative and summative assessment'. Thus assessment *for* learning and *of* learning are not assessments that have to develop independently of each other as they both focus on improved learning.

Possible approaches to assessment

The following assessment approaches were considered for evaluation:

- Standardised testing;
- Common assessment tasks;
- Teacher-generated tasks (including performance assessment);
- Judgment by teacher groups;
- Embedded development and assessment;
- Portfolio construction; and
- Self-assessment.

Other approaches were canvassed in the review of examples. Some were not listed because, even though they are excellent in particular contexts, they were deemed not to be feasible in a school context.

A new picture of assessment and reporting

What emerged from this study is a completely new picture of how senior secondary students' achievement of employability skills, across Australia, might be assessed and reported.

The picture has the following six components:

1. Three suggested assessment approaches (see below), all of which are open to implementation;
2. One preferred approach, selected on the basis of its perceived positive backwash effects on curriculum and teaching/learning;
3. A technique for assessment that involves a two-stage process for arriving at a grade for reporting;
4. Performance levels and facet standards for use in the assessment process;
5. A mechanism for reporting individual student results; and
6. A format for representing results on a student report.

1. Suggested approaches to assessment

After consideration of (a) the five evaluative criteria, (b) the main purposes served by assessment, and (c) seven possible assessment approaches, three of these assessment approaches were found to be fit for the purpose of reporting employability skills achievement. These three assessment approaches are:

- Standardised testing;
- Common assessment tasks; and
- Judgments by groups of teachers from different subjects.

Standardised testing involves students across a jurisdiction (e.g. state or country) of the same year or age sitting for the same test under the same conditions, usually at the same time, with results being reported in a common format. Standardised tests are set and marked by a central agency according to a key (for multiple-choice items) or a commonly applied marking scheme (for constructed responses or short answers). Standardised tests produce reliable results and would enable national comparability but they lack the capacity to assess all employability skills comprehensively (and some of them validly).

An assessment task is a device or constructed situation that creates the opportunity for learners to demonstrate the nature and quality of their learning. It can be long or short, not necessarily written, done in a controlled assessment space or not, completed in a specified time or not, by students working individually or in groups, with or without certain levels of teacher assistance and so on. *Common assessment tasks* are set by a central agency (or group of teaching professionals) and marked by teachers in the students' school according to the same criteria and standards as those used by teachers in other schools. Common assessment tasks would provide a degree of comparability that could be enhanced through moderation, although this would result in increased cost and complexity. The local sites for assessment, albeit where common marking schemes are used, present a reliability challenge. It is possible to develop common assessment tasks for all the employability skills although standardised tests could be an efficient assessment instrument for testing certain facets within the larger task structure.

For the approach referred to as *Teacher-group judgments*, groups of teachers who have taught certain students meet and, using common criteria, make judgments about the achievement

levels of individual students on each of the employability skills. The evidence of achievement on a particular employability skill is gathered from several subject contexts by the teachers of those subjects and also from co-curricular activities where the teachers have interacted with the students. Teachers in the group consider each employability skill in turn for each student and seek to reach agreement about that student’s achievement of that skill. Although consensus judgments have been shown to occur with this method at the school level, a lack of consistency of judgments between schools could be expected and so comparability of reported achievements would be limited. Implementation of this method would require a modest level of funding for teachers’ professional learning.

These three approaches are not the same in terms of where control is situated for setting and marking. Standardised tests, for example, have an external source controlling both setting and marking. Teacher-group judgments have an internal source controlling both setting and marking (subject teachers and teacher groups, respectively). Common assessment tasks have an external source (e.g. assessment agency or teacher professional community) controlling setting and an internal source (teachers) controlling marking, although externally developed marking schemes are used.

Marking	Setting	
	External	Internal
	Centrally devised	School devised
External Centrally organised	Standardised test	
Internal Teacher judgment	Common assessment task	Teacher-group judgment

Analysis of stakeholder perspectives on the preferred options

Each of the three suggested approaches was analysed in terms of its perceived benefits and weaknesses from the perspectives of students and parents, teachers, and employers and further and higher education providers. This analysis would be an appropriate starting point for further consultation with stakeholders, not only for checking whether it accurately portrays the views of stakeholders but also for highlighting that the final selection of one model is ultimately an on-balance judgment.

2. Preferred approach to assessment

In addition to the application of the five evaluative criteria to possible assessment approaches, a further criterion was considered—‘backwash effect’; that is, the influence, positive or negative, of the chosen assessment method on curriculum and the teaching–learning process. Of the three suggested approaches, one of them—common assessment tasks—became the preferred approach. This method meets the specified evaluative criteria at least as well as any other while at the same time being the most likely to have positive backwash effects on teaching and learning. An advantage of the method is that it is likely to result in enhanced achievement of employability skills at the cohort level because all students would benefit from the feedback on their learning that task-based assessment provides. Because of the need to maintain security over items in standardised tests, opportunities for feedback are unlikely to arise. Nor are there opportunities for feedback to students in the case of teacher-group judgment, which occurs after a sustained period of observation. It is the common assessment task option that is capable of providing immediate feedback and therefore most likely to lead to gains in students’ understanding and application of employability skills.

In summary, the common assessment task is regarded as being most likely to lead to enhanced achievement of the full range of employability skills at the cohort level, to provide a sound basis for reporting, and to generate an acceptable degree of national consistency. A variation of this method would have standardised testing of some of the facets within common assessment tasks.

3. Technique for assessment

One of the strengths of the employability skills framework (ACCI & BCA, 2002) lies in its capacity for customisation. Each of the eight employability skills in the framework is elaborated through a set of facets. These facets illustrate how an employability skill might apply to particular work contexts. Because the context being considered for this report is schools and because the target group of young people are those who will complete schooling and enter the labour market, a subset of facets, considered to represent entry-level requirements, was selected for each employability skill. This selection constrained the assessment task facing schools to a manageable one.

Classification of facets

Facets of an employability skill were classified as *foundational* or *applied*, with the former subset deemed to be assessable in the schooling sector. Typically, there are four foundational facets for each employability skill.

Two-stage assessment model

Assessment of a particular employability skill is a two-stage process. The two stages are:

1. The teacher–assessor judges the standard of student work on each foundational facet against descriptions of three *facet standards*. These ‘facet-standard descriptors’ describe what a teacher should see in a student’s work at a minimally acceptable, intermediate, and high standard of achievement.
2. The teacher–assessor decides on the overall level at which the student is performing on that skill by judging the combined results from Stage 1 against descriptions of three *performance levels*. Only the performance level (grade), not the facet standards, is reported.

Options for combining facet standards to give performance levels

Two techniques are recommended. One involves combining scores (3 for high standard, 2 for intermediate standard, and 1 for minimally acceptable standard). This process could be automated to eliminate the potential for arithmetical error. The other technique requires teacher–assessors to make on-balance judgments, say trading off a high standard on one facet with a minimally acceptable standard on another facet.

4. Performance levels

In order to apply the two-stage model described above, performance levels for the employability skills were developed for reporting. Three general performance levels were identified for each employability skill for describing the behaviours that a young person might demonstrate as evidence of achievement of that particular skill. Taking the employability skill ‘Communication’ as an example, these levels are:

Performance level	Grade	Description
3	Advanced achievement	Student independently completes complex tasks in multiple environments.
2	Creditable achievement	Student completes tasks in mainly unstructured environments with minimal teacher help and support.
1	Basic achievement	Student completes routine tasks in structured environments usually with teacher help and support.

The function of the descriptors for performance levels is to assist teachers in assigning an overall grade for a particular employability skill. Where there is no evidence of achievement on one of the employability skills, no grade is assigned to that student for that skill for reporting purposes. Ungraded status on an employability skill does not necessarily mean that the student is completely lacking in the skill (although s/he might be). It might simply mean that the opportunity to demonstrate achievement was not available to the student.

The performance level descriptors that have been developed require verification ‘in the field’.

Facet standards

Prototype standards for the foundational facets were also developed. Three specific standards were identified for each facet, customised to the school context, for judging the quality of a student’s work.

The function of the descriptors for facet standards is to contribute to the consistency of teacher judgments about student achievement on a particular employability skill.

Taking ‘Reading independently’, one of the facets of ‘Communication’, as an example, these standards are:

Facet standard	Description
3	Recognises the value of the structures, features and conventions used by authors for constructing meaning in a range of multi-modal texts
2	Extracts implied meaning in most text modes
1	Extracts literal meaning in everyday or standard written texts

The prototype facet-standard descriptors require endorsement.

5. Mechanism for reporting

Reporting gives, for each student, for each employability skill, a grade of ‘Advanced’, ‘Creditable’, or ‘Basic’, with an ungraded category available for students who have not been able, for one reason or another, to provide evidence of achievement on a particular skill.

The suggested mechanism does not privilege one skill over any other skill. The impression gained from viewing results thus presented is of the variation in the pattern of results between students. An implied advantage is that different profiles would appeal to different employers (it is highly unlikely that all students would perform at Level 3 – ‘Advanced’ – in all skills).

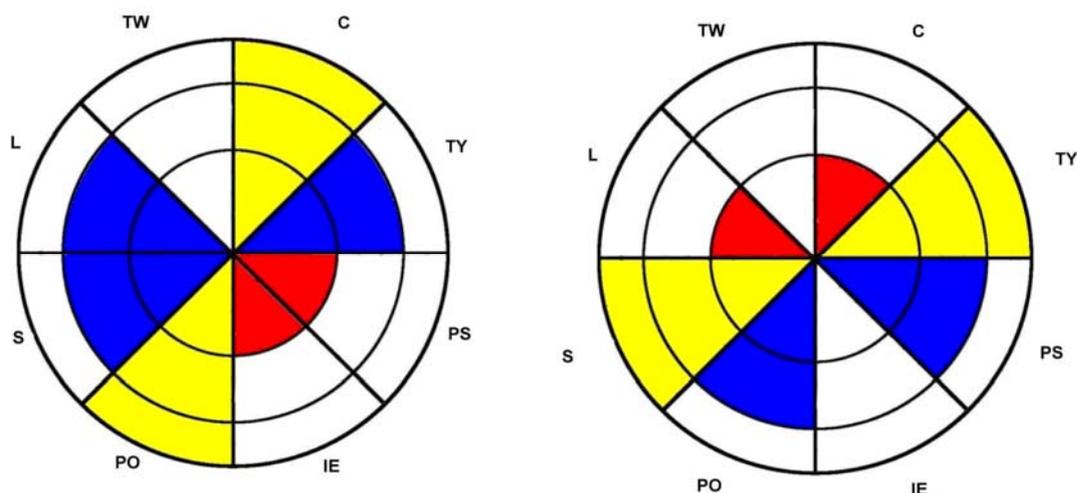
6. Reporting format

Two options are provided for communicating the performance levels attained by a student on each of the employability skills. One is a visual display; the other is simple listing of results. For the former, the suggested format for reporting the results of individual students is a circle with eight segments. In this visual representation, all of the eight skills are in contact with each other (at the centre of the circle). The pattern of results is thus easily stored in the memory of the reader of the report. This would not be the case with vertical bars arranged along an axis or with a simple listing of results

Option 1: Visual display

The three identified performance levels, Basic, Creditable, and Advanced, are represented in a special type of bar graph (the ‘bars’ are segments of a circle like slices of a cake), with a blank for ungraded. One possible design uses three colours, one for each of the three performance levels. A simpler version uses only black and white.

The figure below shows a *profile of results* for two students, with colour coding according to level (three colours). In both figures, a blank segment indicates ‘ungraded’ status.



Colour code

Yellow – Advanced; Blue – Creditable; Red – Basic; No colouring – Ungraded

Legend

TW Teamwork

C Communication

L Learning

TY Technology

S Self-management

PS Problem Solving

PO Planning & Organising

IE Initiative & Enterprise

Option 2: Simple listing

Students 1 and 2 have been awarded the following grades for their achievement in the employability skills:

Employability Skill	Student 1	Student 2
Communication	Advanced	Basic
Technology	Creditable	Advanced
Problem solving	Basic	Creditable
Initiative & Enterprise	Basic	Ungraded
Planning & Organisation	Advanced	Creditable
Self-management	Creditable	Advanced
Learning	Creditable	Basic
Teamwork	Ungraded	Ungraded

Chapter 1 – Introduction

Background to the study

This report documents the findings of a study designed to examine possible approaches to assessing and reporting the achievement of employability skills by senior secondary students.

The objectives of the project are:

- To identify and consider appropriate methods for the assessment and reporting, on a nationally comparable basis, of Year 12 students' achievement in relation to the eight employability skills specified;
- To ensure that the proposed methods for assessment and reporting are valid, reliable, objective, feasible, usable, and nationally comparable; and
- To ensure that the proposed methods for assessment and reporting can evidence the employability skills already embedded in curriculum and pedagogy.

The project, which was undertaken between 1 July 2007 and 31 January 2008, included four activities:

- Identification of performance levels for each employability skill that are appropriate for students completing Year 12 in Australia;
- Examination and description, taking account of existing national and international research, of a range of methods for assessing and reporting on student achievement against the employability skills in the senior secondary years in terms of validity, reliability, objectivity, feasibility and usability;
- Overall evaluation of the most appropriate approach to assessing and reporting on each of the employability skills in senior secondary school students, together with an analysis and discussion of the extent to which the preferred approach would be of benefit to students, parents, schools, employers, and further training/education providers in specified contexts; and
- Report that clearly sets out the findings and is accessible to parents, schools and employers, covering all the areas described above.

An adjunct activity was to determine the current level of employer satisfaction with the eight identified employability skills and how employers assess them, to compare employer satisfaction with teacher/school views of employability skills, and to identify employer suggestions for improving assessment and reporting strategies used with secondary students. Some results from this activity (Costley, Power, Watson, Steele, & Sproats, 2007) are referred to in the main body of the report.

As a precursor to the project, definitions of the eight employability skills, including their facets, were reviewed, specifically to ensure their conceptualisation as assessable constructs.

Outcomes required of the project are:

- Described performance levels, one set per skill;
- Guiding principles for assessment and reporting of employability skills achievement;
- Options for assessing and reporting each of the eight skills;
- Recommendation of an approach that is most appropriate for assessing and reporting each of the skills; and
- Analysis of the extent to which the preferred approach will be accessible to students, parents, teachers and employers.

Each of the required outcomes is arrived at within the main body of the report except for the last, which is an appendix.

Milestones in the development of employability skills in Australia

Australia, in common with many other countries, has a substantial history of developmental work on the definition of generic employability skills and of their implementation in three sectors—schooling, vocational education and training (VET), and higher education.

Quality of Education Review Committee

The influence of economic and technological change in Australia and of increasing competition within the country and from competitors abroad has contributed to the recognition of a need to improve the aggregate skill levels of Australians. The Quality of Education Review Committee (1985: 204–05) was asked to:

... identify strategies that would ensure the attainment of a satisfactory standard by the great majority of students at successive stages of the general curriculum, with particular reference to communications, literacy and numeracy [and] an improved relationship between secondary education and employment and tertiary education opportunities and requirements.

This remit was further elaborated, requiring the committee to provide advice on means for ‘attaining higher basic skill standards’ for primary school students, and for ‘the attainment of appropriate standards relevant to subsequent employment opportunities and improved preparation for tertiary education’ for secondary students. The terms of reference also included mention of the ‘increasingly competitive, including internationally competitive environment’ of Australian industries into which school leavers would move.

The first recommendation called for improved development (particularly at junior primary school level) in five general competences; namely,

1. Acquiring information;
2. Conveying information;
3. Applying logical processes;
4. Performing practical tasks as individuals; and
5. Performing practical tasks as members of a group (Recommendation 1, p. 201).

The Committee clearly believed that these competences should be assessed using criterion-referenced tests and that they should be reported upon every three years (Recommendation 2, p. 201).

The Committee also exhorted the government to ‘sustain its efforts in curriculum development with particular reference to communication skills, mathematics, science, technology, the world of work and Australian studies’ (Recommendation 10, p. 203).

Thus the Committee recognised the importance of both basic and generic skills and recognised the importance of assessing them. Their recommendations became the focus of deliberations of the Finn and Mayer committees.

Finn Review

The Finn Review (Australian Education Council (AEC), 1991) was asked, among other wide ranging terms of reference, to report on ‘appropriate national curriculum principles designed to enable all young people ... to develop key competencies’ (p. 2). These competencies were to be ‘related to a young person’s initial and lifelong employability’ (p. 54). Two issues that have arisen since the Committee’s deliberations were thus foreshadowed; namely, the notion that generic skills (key competencies) are required in emerging work contexts and that such skills contribute to the ongoing employability of individuals. The Finn Committee recognised that lifelong employability depended upon individuals having ‘a strong grounding in generic,

transferable skills' (p. 55). The Committee (AEC, 1991: 58) identified six key areas of competence:

1. Language and communication;
2. Mathematics;
3. Scientific and technological understanding;
4. Cultural understanding;
5. Problem solving; and
6. Personal and interpersonal understanding.

Because the Finn Committee had responsibilities elsewhere, it recommended that an expert group be established to refine the six areas of competence, recommend appropriate levels of achievement for them, and recommend arrangements for their assessment and reporting. The Mayer Committee was established to undertake that work.

Mayer Committee

The Mayer Committee (AEC, 1992) consulted very widely in developing its definition of key competencies. Through these consultations, the committee was urged to adopt a very wide range of skills and attributes, and it was very careful to define what was meant by a key competency and by each of the seven key competencies that were recommended. The Mayer Committee's requirements of a key competency (AEC, 1992: 13) were that it:

- Be essential to preparation for employment;
- Be generic to the kinds of work and work organisation emerging in the range of occupations at entry levels within industry rather than occupation- or industry-specific;
- Equip individuals to participate effectively in a wide range of social settings, including workplaces and adult life more generally;
- Involve the application of knowledge and skill;
- Be able to be learned; and
- Be amenable to credible assessment.

Values and attitudes, therefore, did not conform to the Committee's stated requirements for key competencies. Hence, the Committee was reluctant to include them. The transition of young people from education into the labour market, rather than the needs of more experienced workers, was a particular concern of the Committee, indicated by their reference to 'preparation for employment' and 'entry levels'.

The Committee paid particular attention to assessment of key competencies. It noted the importance of the application of knowledge and skills and was cognisant of the influence of different contexts of application, and so it recommended assessment for each key competency in at least two different contexts.

There was considerable debate about the number of performance levels that should be recognised, even at entry level. The Mayer Committee recommended three performance levels for each key competency. These levels were described (AEC, 1992: 18) thus:

- Performance Level 1 describes the competence needed to undertake activities efficiently and with sufficient self-management to meet the explicit requirements of the activity and to make judgments about the outcome against established criteria.
- Performance Level 2 describes the competence needed to manage activities requiring the selection, application and integration of a number of elements, and to select from established criteria to judge quality of process and outcome.

- Performance Level 3 describes the competence needed to evaluate and reshape processes, to establish and use principles in order to determine appropriate ways of approaching activities, and to establish criteria for judging quality of process and outcomes.

The descriptors above are actually meta-descriptors, which reflect increasing self-directedness and adaptation from Level 1 through Level 2 to Level 3 (the highest level of performance). If they were to be applied as they stand, the *performance* criteria for assessment would be the same as the *overarching* criteria (or levels) for assigning an overall grade. Teacher-assessors (markers of student work) would then be required to look for primary evidence of self-directedness and adaptation, the latter terms being short versions for descriptions of the performance dimensions for Level 3 above. It would be helpful to markers if the overarching criteria were customised to each skill so that the specific skills were being assessed. Had assessment of the key competencies gone ahead without this having been done, the validity of the assessment would have been greatly compromised.

Throughout the mid-1990s, much work was done to implement the key competencies, especially in the schooling and VET sectors. In an extensive review of the pilot project for the key competencies in New South Wales, Ryan (1997) reported that, in schools and in the VET sector, assessment had presented difficulties for teachers and that the performance levels as described in Mayer's (AEC, 1992) report had not been well understood. (Teachers were required to apply standards in the key competencies to existing subject-based assessment tasks). Ryan suggested (1997: 39) that graded samples of student work be provided. Teachers graded three common assessment tasks undertaken by 800 Year 11 students, and Ryan reported that they were able to distinguish five performance levels. This indicated that, with appropriate support, experienced assessors were able to grade student work consistently. Not for the first time in standards-based assessment did it become obvious that standards (or achievement levels) are not merely words on a page: they do not become established overnight. Rather, they 'become alive' when exemplified by student work that purports to meet the standard and when consensus is reached among expert judges that the evidence (student work) does indeed meet the standard (Pitman, O'Brien, & McCollow, 1999).

Concerns were expressed about the implications for teachers' workload if reporting of achievement in the key competencies were to be separate from reporting achievement in specific subjects/disciplines. The workload issue became one element of a sequence of events that halted implementation of the key competencies initiative in the schooling and VET sectors.

Shortly after the publication of the Mayer Committee's report, there was an expectation that the considerable work already undertaken on national curriculum statements and profiles would lead to their implementation. But criticism of these standards in some jurisdictions thwarted national agreement. At about the same time, breakdowns in negotiations over award conditions for teachers resulted in the imposition of work bans, one of which related to the key competencies (see multiple references in Lokan, 1997). And so, despite considerable progress in addressing some very difficult issues in the definition and assessment of generic skills through the key competencies initiative, that initiative stalled.

The Mayer Committee had recognised (AEC, 1992: 9) the necessity for the key competencies to 'be reviewed periodically to ensure that the set appropriately reflects the generic competencies essential for effective participation in the emerging forms of work and work organisation'.

Employability skills for the future

The extensive work described above illuminates several difficulties that were experienced in implementing the key competencies, such as the lack of clarity in definitions of core constructs and the problems associated with assessment. In the decade since the publication of the Mayer Committee's report, the context in which Australian industry operates had become increasingly competitive, and technological change had continued unabated. With this change in operational context came the recognition that key competencies or employability skills needed to be embedded more deeply in Australia's education and training systems.

The Australian Chamber of Commerce and Industry (ACCI) and the Business Council of Australia (BCA) led a project to review and clarify concepts and terminology, review generic employability skills developments in Australia and overseas, and recommend a framework of employability skills for implementation in schools and the VET and higher education sectors (ACCI & BCA, 2002). Australia's other major peak employer body, the Australian Industry Group (AIG), in also recognising the need for higher levels of employability or 'soft' skills among both entry-level and existing workers, recommended an increased 'emphasis on development of employability skills' (AIG, 2004: 50, 103).

The framework that emerged from the ACCI and BCA project included a set of personal attributes and a set of eight key skills, namely:

1. Communication skills that contribute to productive and harmonious relations between employees and customers;
2. Teamwork skills that contribute to productive working relationships and outcomes;
3. Problem-solving skills that contribute to productive outcomes;
4. Initiative and enterprise skills that contribute to innovative outcomes;
5. Planning and organising skills that contribute to long-term and short-term strategic planning;
6. Self-management skills that contribute to employee satisfaction and growth;
7. Learning skills that contribute to ongoing improvement and expansion in employee and company operations and outcomes; and
8. Technology skills that contribute to effective execution of tasks.

Each of the skills was elaborated through descriptions of facets of the skill—aspects of how the skill might be applied in particular contexts. It is important to note that the facets were not designed as a mandatory list of sub-skills or as an exhaustive list of aspects.

The report made this clear in the statement (ACCI & BCA, 2002: xiii) that:

- The mix and priority of these facets would vary from job to job.
- The list of elements is not exhaustive but rather reflects the information provided by the specific enterprises interviewed for this study.
- The list of elements is indicative of the expectations of employers.
- The level of sophistication in the application of the element will depend on the job level and requirements.

As part of the ACCI-BCA project, extensive consultations were undertaken with industry and employer groups (Field, 2001; McLeish, 2002) and, subsequently, with unions, parents, schools, the VET sector, and universities. The list of key skills was accepted as appropriate for employment requirements. Further, the Ministerial Council on Education, Employment, Training and Youth Affairs (MCEETYA) (see below) endorsed the employability skills as part of the suite of capabilities that young people require in making a successful transition from school to work.

Australian Certificate of Education

In May 2006, DEST commissioned ACER to investigate and report on models and implementation arrangements for an Australian Certificate of Education (ACE) for the final years of secondary school. In completing this assignment, Masters, Forster, Matters, & Tognolini (2006) came to a somewhat broader understanding of ‘an ACE’ than the original claim. Since the release of that report, Recommendations 1, 2 and 5 have shaped further research questions (on curriculum essentials and achievement standards) and new policy (on a national standards body), respectively. The formulation of Recommendation 3 (on key capabilities assessment) acknowledged the Employability Skills Framework of ACCI and BCA (2002) and encouraged schools and awarding bodies to ‘develop, assess and report on general skills required for life and work beyond school’ (Masters et al., 2006: vi). Of particular relevance to the current study is Recommendation 3, ‘the introduction of a national test of a small number of these skills: initially, reading literacy/verbal reasoning, mathematical literacy/mathematical reasoning, written English [expression], and ICT literacy’, noting that ‘further work will be required to develop and investigate valid and reliable ways of assessing [and reporting] all eight employability skills’ (Masters et al., 2006: 81).

The issue of assessment and reporting had been identified (Curtis & McKenzie, 2002: 56–59), amongst other issues, as requiring resolution in order to move forward with implementation of the employability skills. While success appears to have been achieved in these matters in the VET sector—albeit after much exploration—similar success appears to have eluded the schooling sector. The current project might contribute to redressing that gap. Table 1, which concludes the chapter, summarises milestones in the history of developing and implementing generic skills and employability skills.

Related initiatives

Generic/employability skills are not the only initiatives that have been proposed for refocusing attention on nationally agreed and appropriate outcomes for education and training. Several other initiatives warrant specific attention, although other related developments (e.g. lifelong learning) are worthy of discussion. In addition, there has been ongoing developmental work in arrangements for senior secondary studies and certification in the various jurisdictions.

National Goals for Schooling

Originally developed by the AEC (1989)¹, the set of ten ‘Common and Agreed National Goals for Schooling’ was revised and, in 1999, endorsed by MCEETYA to become the National Goals for Schooling (MCEETYA, 1999)². The Adelaide Declaration expressed aspirations for student achievement of basic (Goal 2.2) and generic skills (Goals 1.1 and 1.3), and access to both general and vocational education (Goals 1.5, 2.3 and 3.6). Goals 1.5 and 2.3, respectively, express the following aims for students.

[Possession of] employment related skills and an understanding of the work environment, career options and pathways as a foundation for, and positive attitudes towards, vocational education and training, further education, employment and life-long learning; and

[Participation in] programs of vocational learning during the compulsory years and [having] had access to vocational education and training programs as part of their senior secondary studies.

¹ The Hobart Declaration

² The Adelaide Declaration

Subsequent MCEETYA resolutions have resulted in further refinement of these national goals; for example, in 2003 MCEETYA resolved that:

There is in-principle agreement that the eight employability skills groupings do include skills that young people require for successful transition from school, but they are not prescriptive and need to be considered as part of the broader set of generic skills, capabilities and understandings that are fostered by the National Goals of Schooling. In this context it is agreed that the Employability Skills Framework be referred to the AESOC³ project on developing Nationally Consistent Curriculum Outcomes. Implementation of employability skills (including assessment and reporting) is a matter for individual jurisdictions and will take account of local policies and practices.

The National Goals for Schooling have influenced the development of many other policies designed to promote successful transitions from school. Two of these are described below.

Enterprise and Vocational Education

The National Goals for Schooling require that students have access to vocational learning in the compulsory years of schooling and to vocational education and training in the senior secondary years.

The goal of increasing access to vocational education for senior secondary students has been very successful. Between 1996 and 2004, the number of students participating in school-based VET subjects grew from fewer than 20% to almost 50% of the cohort (MCEETYA, 2004). Access to structured workplace learning opportunities has not kept pace with the growing demand for VET study and there is considerable variation between jurisdictions in the intensity of VET offerings (Lamb & Vickers, 2006). The status and quality of VET offerings in schools are rising as more schools offer full qualifications rather than subjects or units and these offerings comply more closely with the Australian Quality Training Framework (AQTF) requirements. Despite the growth in school-based VET participation, the report *Learning to work* recommended that:

... enterprise and employability skills be made a higher priority and developed through a range of strategies across the curriculum in addition to the VET in Schools pathway, to maximise the effectiveness of vocational education in preparing students for post-school options.

(Australian Parliament. House of Representatives Standing Committee on Education and Training, 2004: xxx–xxxii).

The Adelaide Declaration also requires exposure to enterprise skills (Goal 2.4), and several major projects have been undertaken to foster these skills. The (then) Department of Education, Science and Training⁴ (DEST) provided considerable support for enterprise education through the establishment of the Enterprise and Career Education Foundation (ECEF). The key role of ECEF is to encourage linkages between schools, communities and industry, through these to promote access to structured workplace learning, to conduct research into effective practice, and to evaluate practices.

Enterprise skills are included in the Employability Skills Framework (ESF) under ‘Initiative and enterprise skills’.

Australian Blueprint for Career Development

Closely related to enterprise education is career education. Through career education, students are expected to learn about the nature of work and to evaluate their interests when making decisions about their future careers. A key initiative in career education was the

³ Australian Education Systems Officials Committee

⁴ Since the 2007 federal election, the Department of Education, Employment and Work Relations (DEEWR)

development of the Australian Blueprint for Career Development (ABCD) (Haines, Scott, & Lincoln, 2006).

The ABCD has two main components: a list of career competencies that people of all ages require in order to manage their lives, learning and work; and a set of processes for planning, implementing and evaluating career decisions. Some career competencies are similar to some skills in the ESF; for example, 'Building positive relationships' (ABCD) and 'Teamwork' (ESF); 'Lifelong learning' (ABCD) and 'Learning' (ESF). Two of the life-stages recognised in the Blueprint are directly relevant to secondary students as they plan their future after exiting from school.

Emergence of employability skills in vocational education and higher education

While there has been a desire to achieve a common approach to employability skills across the three education sectors in Australia (Allen Consulting Group & National Centre for Vocational Education Research, 2004), the contexts of each have led to diverse approaches in them. It is possible, however, that developments in the VET and higher education sectors can inform employability skills assessment and reporting in the schooling sector. See Table 1 at the end of this chapter for a summary of key generic skills developments in the three education sectors.

The VET sector

The key competencies were incorporated in training packages in the VET sector. Typically, however, they were represented as a table listing the seven key competencies with three performance levels and an indication of the level expected to be achieved for each of these competencies. In some training packages, modifications were made to evidence and assessment guides, but it appears that these changes had little impact on the delivery of training. A significant exception was the initiative undertaken at Torrens Valley Institute of TAFE⁵. Flagged as a good practice exemplar in a review that could find few such examples (Clayton, Blom, Meyers, & Bateman, 2003), the model embedded the development and assessment of key competencies (and subsequently employability skills) in the teaching and assessment of units and led to the separate reporting of key competencies achievement.

An extensive period of exploration of options occurred following the publication of the ACCI-BCA (2002) report. One of the reports to emerge during this period (Ratio Pty Ltd & Down, 2003) canvassed eight assessment options, but there was reluctance to implement any of them. One of the options was a 'mapping' exercise in which evidence of the various employability skills would be identified in assessment requirements and evidence guides of training packages. It was expected that this approach would have little impact on the development of these skills through the delivery of training. Other options were met with a lack of enthusiasm due to the overheads in their implementation. For example, professional learning was expected to be quite expensive and the rewriting of training packages around employability skills was rejected due to its resource-intensive nature.

One of the recommendations of the House of Representatives Standing Committee on Education and Training (2004) was the creation of Certificate I level qualifications in employability skills. Three such certificates were developed, but they were not included in training packages. Such qualifications may serve a useful purpose for young people who face major challenges in finding work or training, but a solution was required that integrated the development and assessment of employability skills within existing industry-recognised qualifications. Recently, the National Quality Council (NQC) (2007) endorsed an approach originally recommended by the Allen Consulting Group (2006). In this approach training

⁵ Technical and Further Education

packages are being revised to include reference to facets of the employability skills in various components of the packages, specifically as elements of competency, in assessment requirements and evidence guides. Employability skills are not separately reported just as other component competencies are not specifically reported. Rather, employability skills achievement is reported using brief descriptive outcome statements at the qualification level. Thus, if an individual fulfils the requirements of a qualification, it is deemed that s/he has achieved the component competencies, including the employability skills. An analogue of this approach for the schooling sector would be to integrate employability skills into syllabus statements and assessment guides for senior secondary subjects.

The VET model is not directly applicable to the schooling sector, as there is no structure in which particular subjects contribute to differentiated, vocationally specific qualifications. Rather, a single qualification (or in one jurisdiction, one of two qualifications) can be achieved by pursuing a very wide range of subjects. If this model were pursued in the schools sector, employability skills would have to be recognised within all subjects because of the diversity of subjects that can contribute to the senior secondary qualification.

The higher education sector

Generic skills, known by a diversity of labels, have been abroad in the higher education sector in Australia in a formal sense since about 1990 (Milne, 2000). Crittenden (1997) argued that certain generic skills have always been the hallmark of university education. The report *Achieving Quality* (Higher Education Council & National Board of Employment Education and Training, 1992), written shortly after the Dawkins reforms (Dawkins, 1988: 287) that led to the 'unified national system' of higher education and a rapid expansion in the number of institutions identified as universities, flagged the need to monitor the outcomes of higher education. The report placed generic skills on the higher education agenda. The acquisition of generic skills was regarded as a potential indicator of higher education quality, and generic skills have indeed been used in quality assurance processes and in the allocation of funding. (See, for example, Access Economics, 2005).

Many universities embraced generic skills schemes, but there was resistance to a common set of skills. In part this might be attributed to universities collectively wanting to differentiate themselves from the other education sectors and also, individually, to differentiate themselves from each other. Thus, each university developed unique statements of 'graduate attributes' and they used a variety of labels for them. It is possible to map the schemes adopted by various universities against other generic skills schemes; for example, key competencies or employability skills. Typically, there is great commonality between them. Communication, teamwork and problem solving feature universally in the schemes and are included in the ESF. Other skills such as 'Critical thinking' and 'Using professional knowledge' are frequently found in university schemes.

The Australian Technology Network (ATN) universities collaborated in a project to identify a common set of descriptors for that group (Bowden, Hart, King, Trigwell, & Watts, 2000). Their report recommended, however, that the selection of particular attributes and the assessment of these are matters that should be left to discipline groups within universities.

Through the Business Higher Education Round Table (BHERT) and the Business, Industry and Higher Education Collaboration Council (BIHECC) there has been a continuing dialogue between business and universities on the promotion of generic skills (BHERT, 2003). Universities have trialled the development of student portfolios as vehicles by which graduates can consolidate evidence of their achievement of graduate attributes, although concerns have been expressed recently about the cost to universities of maintaining electronic (digital) portfolios and about an implied endorsement of their content if they are maintained on university websites (Precision Consultancy, 2007).

There may be lessons from the higher education sector and other sources (Curyer, 2006), on the feasibility of portfolios for the schooling sector. The element of comparative advantage for differentiated generic skills schemes in the higher education sector is not likely to inform the deployment of the ESF in schools.

Target groups of students

Figure 1 shows the dominant paths of students from compulsory education into the labour market. The patterned areas in the figure represent the points at which employability skills that might have been developed could be assessed (and reported on).

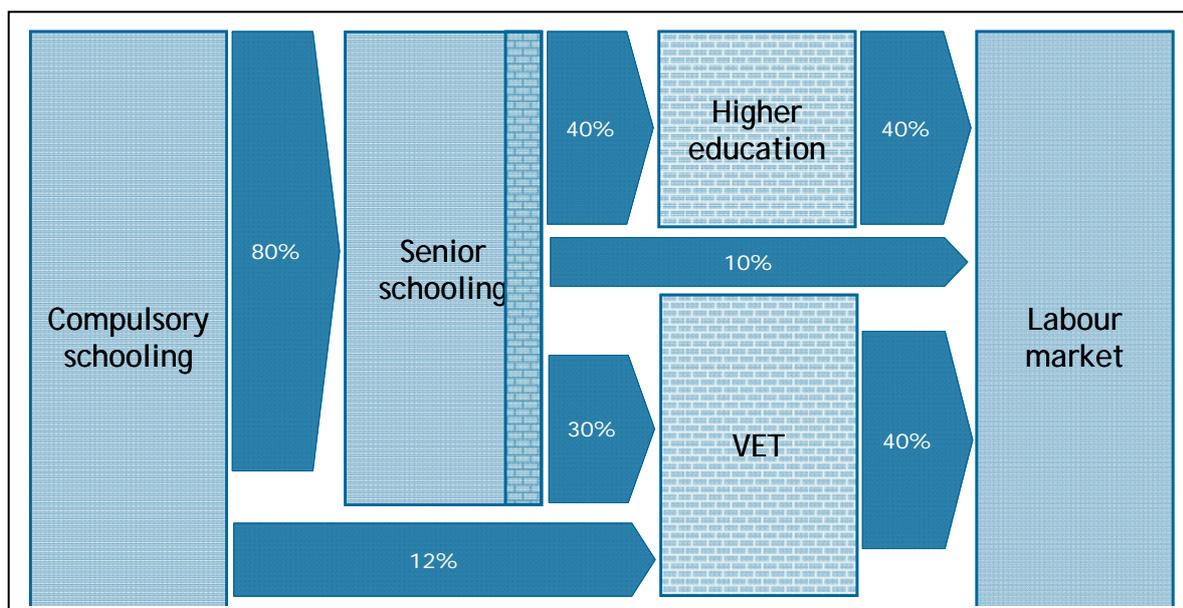


Figure 1: Opportunities for the development, assessment and reporting of employability skills for students on various paths from compulsory schooling into the labour market⁶

Sources: ABS (2007), Curtis, D. D. (forthcoming).

Students who complete secondary schooling and seek to move directly into the labour market are the target group for the current study because, for those students, any proposed assessment and reporting activities would be their only formal exposure to these constructs (Employability Skills).

Students who engage with VET or higher education, whether they complete senior schooling or not, would have further opportunities to develop their employability skills through their post-school studies. Students who leave school before completing senior secondary schooling and who do not engage with VET would not have their employability skills assessed and reported. Consideration might be given to introducing these skills in early secondary schooling, perhaps in conjunction with career education (Haines et al., 2006; McMahon, Patton, & Tatham, 2003). Some vehicles exist for this integration (e.g. the proposed Personal Learning Plan of the *future* SACE initiative and the Learning Account in the new QCE).

Conclusion

This chapter concludes with a tabular summary of milestones in the history of the development and implementation of generic skills/employability skills in Australia.

⁶ Because of attrition and multiple enrolments, the percentages displayed are approximations and may not sum to 100% at all stages between schooling and the labour market.

Table 1: Summary of generic/employability skills developments in Australia, by sector

Sector	Schooling	VET	Higher education
Milestone			
Articulation and definition	Identification of the need for higher-level skills (Quality of Education Review Committee, 1985)		
	Identification of generic skills (AEC, 1991) Identification of key competencies; proposals for assessment (AEC 1992)		Institutional developments beginning early 1990s (e.g. Milne, 2000; Bowden et al., 2000; University of SA, 2000))
Implementation of key competencies	Key competencies in schools (MCEETYA Transition from Schools Taskforce. Working Group on KCs, 1996) Review of implementation trials (Ryan, 1997) Factors that militated against implementation of the key competencies (see notes within chapters on statements and profiles in various states, Lokan, 1997) (National Industry Education Forum, 2000)	Early versions of training package guide for developers included advice on inclusion of key competencies in unit descriptions (ANTA, 2004) Review of assessment models for incorporating employability skills in training packages (Ratio Pty Ltd & Down, 2003) Development of certificates in employability skills in three industries (Unpublished report, Mitchell and Associates)	Key competencies were not implemented, although many universities did respond to the 'Achieving quality' report (Higher Education Council & National Board of Employment Education and Training, 1992) Crittenden (1997) argued that universities had traditionally focused on 'generic skills'.
Assessment and reporting	Group-teacher judgment (McCurry & Bryce, 1997, 2000)		External assessment (ACER 2000, 2001)
	Embedded VET in Queensland senior subjects (1994)		
ES definition and initial validation	Background research including consultations and validation (ACCI & BCA, 2002; Curtis & McKenzie, 2002; Field, 2001; McLeish, 2002)		
Implementation		Official endorsement, following extensive background work (National Training Quality Council, 2005)	Promotion of generic skills in higher education (Business Higher Education Round Table, 2003; Hager, Holland, & Beckett, 2002)
		Recommendation of certificate in generic employability skills (Australian Parliament. House of Representatives Standing Committee on Education and Training, 2004)	
Focusing on development and assessment	The Universal Recognition project (Allen Consulting Group, 2004; Allen Consulting Group & National Centre for Vocational Education Research, 2004)		
	Mapping employability skills against curriculum and assessment standards documents (ACACA, 2003) Relating employability skills to other initiatives (Erebus Consulting Partners, 2003)	Canvassing assessment options (Australian National Training Authority, 2002, 2003, 2004; Ratio Pty Ltd & Down, 2003) A recommendation on assessment (Allen Consulting Group, 2006)	

Chapter 2 – Fundamental Concepts and Issues

Approaches to the assessment and reporting of employability skills must be appraised in terms of technical criteria and curriculum criteria. Technical requirements include validity, reliability, objectivity, and some additional characteristics. Curriculum requirements include feasibility and usability. Both sets of criteria depend upon establishing a purpose—or perhaps several purposes—for assessing and reporting achievement. Chapter 2 discusses the purpose(s) for the assessment and reporting of achievement in employability skills as well as the concepts introduced as criteria for the evaluation of possible approaches to that assessment and reporting. The previous chapter documented milestones in the development of employability skills in Australia. This chapter concludes with a discussion of the concepts of generic, work-related and employability skills.

Criteria for evaluating approaches to assessment and reporting

A requirement of the project was the evaluation of potential methods for assessing and reporting employability skills against five criteria:

1. Validity
2. Reliability
3. Objectivity
4. Feasibility
5. Usability.

The meanings of these and of (other) closely related concepts and their implications for the evaluation of possible assessment and reporting approaches are elaborated in this section.

Validity and related constructs

Messick and three leading research associations in the United States, the American Educational Research Association (AERA), American Psychological Association (APA), and the National Council on Measurement in Education (NCME) assert that the validity of an assessment is the extent to which evidence and theory support judgments and interpretations, including the social consequences that follow those judgments and interpretations, that are made from test scores (AERA, APA, & NCME, 1999; Messick, 1995). Messick (1995) argues that a range of sources of evidence has to be assembled in order to establish the validity of an assessment. Sources include evidence about the target construct of the assessment, the generalisability of judgments made on the assessment, and the consequences for individuals of the judgments.

Messick (1998) identifies two key threats to the validity of judgments: construct under-representation and construct-irrelevant variance. There are three implications for this current study of these two threats: first, that the target constructs for the assessment (eight employability skills) must each be fully and clearly described; second, that assessment instruments (tasks or tests, say) must address all aspects of the constructs; and, third, that the assessment instruments should not be assessing other constructs that are unrelated to the target. These requirements are non-trivial, because the target skills do not exist as isolated constructs; each is a skill that embodies certain knowledge and, critically, each exists only in a context. For example, ‘Communication’ is a set of expressive and receptive skills, about which a competent communicator is knowledgeable, and which can be applied in a range of contexts. These contexts have their own bodies of knowledge and their own particular facets of the skill. If communicative competence is being assessed through a history essay, the judgment about the learner’s communicative competence must be disentangled from her declarative knowledge of history. History, however, has its own forms for the presentation of

evidence and for constructing an argument. It is almost certain that a single activity, such as a history essay, will not provide sufficient scope to assess all facets of communicative competence and that a judgment of that target construct will need to be made over several tasks.

The social consequences of judgments for individuals constitute an element of validity and they may be crucial for individuals. Employers may use employability skills reports for recruitment selection and this use gives the assessment a high-stakes status. It is critical, therefore, for selection decisions to be supported (complemented) by the assessment and reporting methods, that the assessments must be informative of an individual's target abilities, free from bias, as well as fulfilling other criteria, discussed below.

Authenticity has been raised as an attribute of good assessment. Authenticity may be an element of validity in that for an assessment task to be valid it should have some predictive power for future performances. An assessment task that reflects likely future contexts possesses both predictive power and authenticity. For Wiggins (1998), the features that define an authentic task are that it is realistic, requires judgment and innovation, and requires action by the learner. Although it is reasonable to expect these features of an assessment task, the assertion of authenticity remains problematic: Authenticity is a matter of degree rather than a matter of the mere presence or absence of some characteristic. The extent to which a task reflects a real-world situation could be judged on several dimensions, and the demands of real-world tasks are themselves quite variable, so the requirements for judgment and innovation would differ markedly across tasks—and, indeed, across setters and assessors of the tasks, depending on their experiences of the wide world beyond the classroom.

School-learning has been criticised as not providing an adequate preparation for 'the real world of work'. Resnick (1987) for example, contrasts the individualistic, abstract, symbolic, and generalised character of school-learning with the collaborative, applied, contextualised, and situational nature of workplace activities. Whether these criticisms of school-learning remain accurate is moot. Certainly, there appears to have been a substantial increase in collaborative work within schools, and curriculum offerings have been broadened to include many more applied subjects than was the case two decades ago. The rapid growth in enrolment in vocational subjects attests to this broadening (MCEETYA, 2004). The capacity of schools, however, to make learning situational is limited. Schools are required to provide a general education for their students whose potential post-school pathways would be infinite in number and whose experienced work contexts of great diversity. Attempts to contextualise and situate school-learning have the potential to undermine the generality required of schooling.

Reliability and related constructs

The consistency of scores or grades resulting from the assessment of employability skills, like the consistency of scores or grades resulting from any assessment, can be estimated using various methods including test-retest reliability. Comments on the reliability of the various possibilities for assessing and reporting employability skills must be preceded by an examination of potential sources of variability in scores or grades.

The target source of variance is differences between individual students' performance on each of the employability skills. A variety of confounding sources are also identifiable. Measurement error in judgment and scoring is an obvious source of variance. It reduces the reliability and precision of measurement but it does not introduce bias. Other sources of variance including differences in teachers' perceptions of a given piece of work and the opportunities afforded by various tasks and contexts for the development and demonstration of skills have the potential to obscure true variance in ability. These latter sources have the

potential to introduce bias and therefore to compromise the fairness of the assessment. Each of these potential sources needs to be identified for each potential method of assessment.

For high-stakes purposes, adequate reliability of a score is not a sufficient criterion. If learners are to be compared by employers on the basis of their reported achievements on each of the employability skills, the precision of the measure must be adequate to support the granularity of decisions that will be made. This has implications for the assessment methods, performance levels identified, and reporting formats that are used. The reporting formats must reflect the precision of the measures that are derived from the assessments of employability skills. Although precision is normally applied to quantitative summaries of assessments, it has an analogue in qualitative judgments, for which the discernment of the judgments must be able to deliver an overall assessment result that supports the required discrimination among candidates.

Moss (1994) challenges heavy reliance on reliability as a criterion that, she argues, privileges standardised assessment by assuming that reliability within or between assessments foreshadows future performances. She argues that the attainment of adequate levels of consistency in teacher judgments has been demonstrated in situations where common tasks are used and teachers are trained to look for particular forms of evidence. Moss suggests that teacher judgments of student performance should include an interpretation of observed performances for future applications. The key issue in balancing the requirements of validity and reliability is establishing the trustworthiness of the score or grade that is reported.

McGaw (2006) similarly argues for the 'fitness for purpose' of assessment, noting that the consequences of reported achievement for individuals have to be considered.

Compromises are required between attaining high reliability and high levels of validity in assessment. Some forms of standardised assessment can yield highly reliable and precise estimates of performance, but may not adequately reflect the scope of the target construct. This is the basis of criticisms of standardised tests in some domains. For example, it may be possible, using a multiple-choice test, to assess with great precision a learner's knowledge of component skills of communication. Such tests however, may not enable the assessment of the candidate's ability to apply that knowledge in novel contexts. On the other hand, assessing the application of these employability skills in real contexts may add to the authenticity and validity of the assessment, but at the cost of reliability and precision. Judgments about an appropriate balance between high levels of reliability and precision and high levels of authenticity and validity will be required, and the balance may differ when different assessment purposes are invoked.

Objectivity

Objectivity in testing arises in at least two ways. First, and conventionally, objectivity is assured by creating tests and test items so that candidates' responses can be scored unequivocally and consistently by anyone who marks the test. This is assured most easily by using structured response (for example, multiple-choice) formats for test items. Alternatively, constructed responses can be used provided the rules for scoring those responses are clear and comprehensive (Choppin, 1997). Second, however, Sadler (1986) argues that subjectivity or objectivity is not a property only of the final act of judging or scoring a piece of work. The decision to test certain characteristics, the choice of the assessment method, and the criteria selected are judgments that precede scoring. That is, in any assessment, a series of judgments is made about what to assess, how to assess it, under what conditions, against what criteria or standards, and, finally, about the quality of the work against agreed standards. In large-scale testing, expert panels are established and processes developed to ensure consistency in the judgments that are made. While all students participate in the chosen assessment methods, students are not necessarily affected equally by those choices. Thus, objectivity may be more

ephemeral than is commonly believed and so it might be the case that other criteria, such as fairness, should carry more weight in the selection of assessment methods, especially for complex skills.

Where open-ended responses (e.g. essays) are assessed, variability in judgment by different raters is constrained by the use of scoring schemes, including rubrics. These are less prescriptive than the rules that might be used to score simpler constructed responses and some post-test checking is required to ensure that objectivity has been achieved. This is done by designing marking arrangements so that at least a sample of student responses is marked by multiple assessors. The marks allocated by different markers to different questions on student scripts are then analysed to ascertain whether there are differences in the severity of markers or the difficulty of questions and whether there are interactions between markers and questions. Where differences are found, marks can be adjusted to ensure that students are treated fairly in the assessment, and feedback to markers can be used to achieve greater consistency in future assessment rounds. This approach is also useful in accommodating the possible effects of rater and question variability so that students who answered different combinations of test items can be compared and graded fairly. In relation to performance assessment, Shavelson, Gao and Baxter (1993) found that, with the support of scoring rubrics, assessor variability is acceptable but that different tasks contribute substantially to score variability. Attention needs to be paid to the objectivity and fairness of those who assess student work and of what tasks are undertaken for assessment.

Feasibility

Feasibility means capable of being done, with the connotation of convenience and practicability in the doing. While many things are doable, fewer are feasible.

The feasibility of assessments of employability skills on a nationally comparable basis requires that all aspects of an assessment meet the criteria of cost- and time-effectiveness and that the benefits flowing from the assessment justify the costs and time required. These aspects include development and validation of assessment instruments, administration of the assessment, gathering students' responses, marking responses, assigning overall grades, recording and analysing student grades, and reporting and maintaining records of student achievement. For standardised assessments, Murray (2003) estimated a practical time limit of 90 minutes for the student testing component. Having students do a test (or assessment task) is only one component of an assessment regime, but greater time might be justified if multiple purposes were to be served by an assessment activity; for example, if employability skills assessment were integrated with existing assessment activities.

The interests of a range of stakeholders need to be considered. State and territory education systems (government and non-government), individual schools, teachers and students are all stakeholders with direct involvement in the assessment of employability skills. For these individuals and groups, the time and resources committed to the assessment must be perceived to be worthwhile. Other parties, notably parents, potential employers and post-school education providers, have an interest in the testing and its outcomes. A net benefit for each stakeholder group needs to be demonstrated, and the benefit clearly must be higher for those assessment and reporting methods that are more resource-intensive. Direct benefits for students and employers include the existence of a report on employability skills achievement. An indirect benefit may be that the act of having these skills assessed signals their importance and should lead to higher levels of employability skills in the student cohort.

The feasibility of assessment and reporting methods may be influenced by a 'backwash' effect—the extent to which the methods influence curriculum content and practices. A positive backwash effect might justify assessment and reporting methods that require greater effort and time than others while a negative effect would militate against those methods.

Usability

The usability of assessment and reporting methods involves the capacity of the assessment and reporting system to be informative to stakeholders in meeting their diverse needs.

The usability of a given approach to assessment and reporting of employability skills depends on how practicable and desirable that approach is. An approach will be regarded as practicable if it works and imposes a justifiable yet limited load upon participants and yields valuable information to stakeholders. Reports of employability skills achievement will be desirable if they are informative and accessible to their intended audiences, if they add value to other available information, and if they are credible. Their credibility will depend upon the agencies who warrant the reports and who attest to the veracity of the reported results.

Learners require feedback about their achievements. This feedback can be summative indicating a judgment of their perceived status on the employability skills; for example, whether they have met a particular benchmark. Feedback may also be formative in that it should indicate to students who have achieved at a particular level what they might do to attain a higher level. Good assessment can provide both forms of feedback. Parents may have similar interests to learners. Indeed parents may want to know in what ways they could support their children in enhancing their employability skills.

Employers have particular interests. They are likely to want succinct summary reports plus access to more detailed information to corroborate and elaborate the summary information. Similarly, post-school education and training providers may be interested in the levels of employability skills of commencing learners in order to tailor their programs to focus upon those skills that most require development.

Teachers and schools may be interested to examine the achievement of employability skills as outcomes of their programs in order to review their activities and to improve their effectiveness.

For reports to be useful to external stakeholders, particularly employers and post-school education and training providers, the reports must have credibility. Credibility will be enhanced if the reports are based upon sound assessment regimes. Credibility of reports will be enhanced if they are issued under the auspices of bodies that can attest to the veracity of the information that is reported. These bodies could be schools or state and territory curriculum and assessment bodies.

Roles and purposes of assessment

Purposes of assessment

Assessment is deemed to have three broad sets of purposes: to promote learning; to measure individual achievement; and to evaluate programs (Airasian, 1994; Pellegrino, Chudowsky & Glaser, 2001). Researchers in the UK and the US refer to assessment *of* and *for* learning (Assessment Reform Group, 1999) and assessment *as* learning (Earl, 2005) while in Australia Forster (2005) refers to assessment for learning and reporting for learning. Assessment of learning is often used to describe assessment that leads to reports of achievement or performance. It is usually associated with the measurement of individual achievement. Assessment for learning is usually associated with the promotion of learning. It describes assessment that is used to inform instruction and therefore to direct or promote learning. Assessment as learning describes the creation of assessment tasks designed to engage students in challenging and meaningful activities through which learning will occur; that is, to integrate learning and assessment. This approach is consistent with the promotion of learning objective. Promoting learning and the greater achievement of employability skills can be driven by assessment practices that focus the attention of teachers and learners onto these constructs. Such attention should lead to higher levels of achievement in the cohorts of

students who experience this attention. While assessment classified as being *of, for, or as* learning does not include explicitly the program evaluation purpose, the aggregation of data on individual achievement through assessment of learning is the most effective way of evaluating programs or systems.

The abovementioned categories of assessment are not necessarily exclusive. Good assessment regimes can be designed to promote learning and to report on that learning (Stobart, 2004). It is unlikely, however, that a single assessment act will promote learning, unless that act is such a high-stakes event that instruction is focused on preparing students for it. It seems very unlikely in the senior secondary context that employability skills will be accorded such high-stakes status. Therefore, an assessment regime is indicated in which a series of learning and assessment activities occurs, through which evidence of performance is used to inform subsequent learning and that accumulates to provide a basis for reported achievement.

The measurement of individual achievement is required if credible reports of employability skills performances are to be produced.

The signalling function of assessment

Assessment signals to teachers, learners, parents and other stakeholders what is regarded as important in curriculum. What is not assessed may be relegated to the margins of the school curriculum. Assessing students' achievement of employability skills raises the status of these skills in the curriculum. Assessing and reporting individual achievement will inform students, teachers, parents and others of individual student achievement and the data can be aggregated, using appropriate methods, to provide information about system-level performance on this important set of indicators.

Assessing students at or near the end of their school education will provide summary data on individual, and potentially system, performance. Such assessment should be a driver for activities embedded in school curricula designed to enhance students' acquisition of employability skills. The development of a national set of instruments for assessing employability skills might lead to greater attention being paid to the development of those skills. Specific attention to the development of students' employability skills via their experiences in existing curricula is likely to enhance the achievement of those skills and the availability of these skills to commerce and industry.

Implications for assessment and reporting of employability skills

The main purpose of the presentation of assessment models in this report is to provide a basis for reporting individual achievement of employability skills. Therefore, in the evaluation of assessment options, those options that are able to support reports of student achievement will be preferred. However, the signalling function of assessment is important. Methods that lead to greater attention being paid to the teaching of facets of employability skills are likely to result in a situation where more students achieve higher levels of these skills, so the likely backwash effects of assessment options need to be evaluated.

Notions of generic, work-related and employability skills

There is a widespread and enduring interest in conceptualising generic and work-related skills in Australia and overseas because of a conviction that the way to help more students build better futures is to develop in them skills that are readily adaptable to the changing social, cultural and technological worlds in which they live. Not narrowly context-specific, these skills emanate from a well-designed and well-delivered curriculum, and are enhanced through extra-curricular activities and life experiences.

Generic, work-related skills have been called different things in different countries as summarised below (see Appendix 1 for a more detailed coverage).

United Kingdom – Core skills, key skills, common skills

New Zealand – Essential skills

Australia – Key competencies, core skills, transdisciplinary skills, repertoires of practice, cross-curriculum skills

Canada – Employability skills

United States – Basic skills, necessary skills, workplace know-how

Singapore – Critical enabling skills

France – Cross-field skills (*Compétences transversales*)

Germany – Key skills (*Schlüssel qualificationen*)

Switzerland – Transdisciplinary goals

Denmark – Process-independent skills.

Because of their nature, vocational or occupational skills are more likely than not to change over time as the demands of work evolve. Work-related skills such as familiarity with information technology and competence in Asian languages reflect the requirements of the late 20th century and would not have had cachet 100 years ago. The desirability of employability skills such as teamwork and the capacity for lifelong learning has not been made explicit until recent times; it might transpire that the predominant culture of the late 21st century deems other skills and attributes to be the desirable ones.

Burke and Eraut (1989) refer to the concept of generic competence as ‘a generalised ability or capacity developed over time from many different and diverse experiences’. This definition sees the performance of non-routine specific tasks as dependent on both general skills and specific knowledge of the context and relevant content. As such, this conceptualisation accords well with other independent work on generalised learning (Annett, 1989; National Curriculum Council, 1990). This conceptualisation is generally supported by research in developmental psychology in terms of the acquisition of disciplines and ‘high literacy curricula’ (especially in literature and mathematics) surveyed by Resnick (1987).

In his report to Britain’s National Council on Vocational Qualifications, Jessup (1990) used the term primary core skills as being synonymous with Burke and Eraut’s general competences. He defined primary core skills as those that underpin almost all performances. He saw, as the essential point about these facets of competence, the extent to which the skills are common to behaviour in different areas and contexts. He further assumed that the acquisition of a core skill in some area of competence and context offers the potential of generalisation or transfer to other areas and contexts which employ the same skill ... and, in doing so, captures the two vital aspects of a generic skill.

It is the concept of being generally applicable that distinguishes a generic skill from a core skill. To describe a skill as ‘generally applicable’ is to say that two people, working in different contexts, could exhibit the same skill: the skill is common to the two contexts and people. Authors such as Anderson, Reder and Simon (1996) go so far as to argue that some skills are transferable, meaning that either of these two people, having developed the skill in one context, could transfer it to another context (where the other context might be an unfamiliar one). Anderson et al. assert that the degree of transfer depends on factors such as the number of shared symbolic components, where attention is directed during learning, and the relation of the material originally learnt to the transfer material. Questions about the specificity of content and context remain in the academic debate about transferability.

While many of the terms are used interchangeably, they are not the same. Whereas core skills, sometimes referred to as key competencies, are a subset of generic skills and often subject to the dictates of social and political forces, generic skills are intrinsic to the nature of schooling. Figure 2 shows the ‘interconnectedness’ of some of these categories of skills. While the figure reflects the element of commonality to all the definitions, the size of its components does not reflect the relative worth or the extent of the commonality. There is a hierarchy—generic skills are the overarching skills and all others are subsets and/or intersections (Pitman, Matters, & Nuyen, 1996).

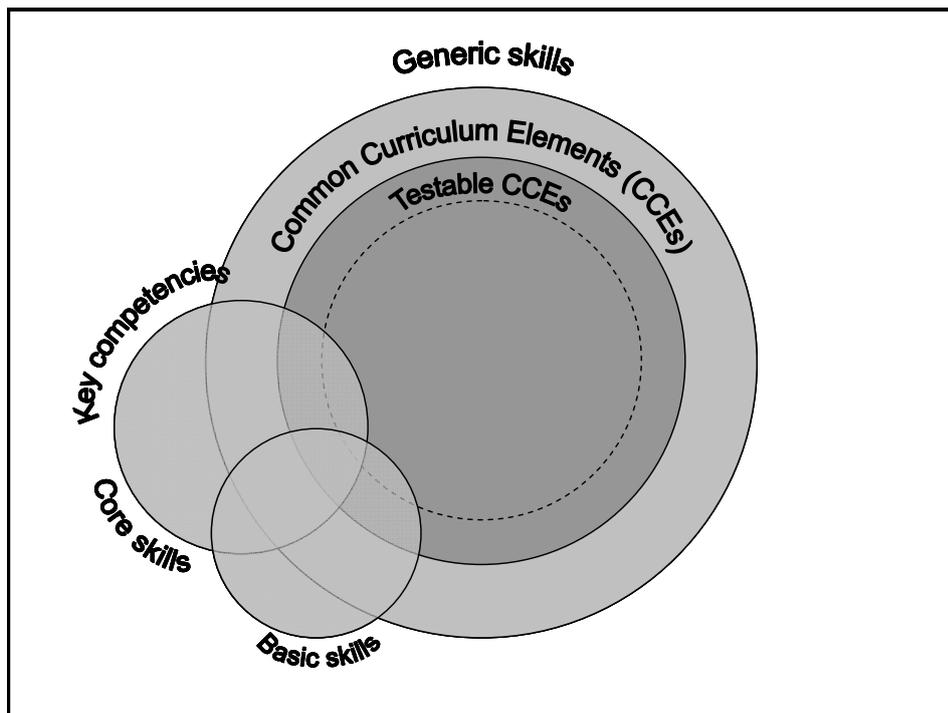


Figure 2: Diagrammatic representation of overlapping sets of skills

The constructs for generic and work-related skills that have been developed in the past 20 years claim to be generic in nature, and they make no reference to the curriculum areas of general education or to subject-specific skills or technical skills. The Mayer Committee’s ‘Using mathematical ideas and techniques’ obviously echoes the subject Mathematics but the sense of it is the *general application* of mathematical ideas and techniques to contexts other than the context in which they were learnt.

In going beyond the cognitive skills into the practical and personal realms, the different work-related skills can be understood and classified (McCurry & Bryce, 2007) according to the following categories.

- Basic, core, key or foundational skills (communication and numeracy);
- Domains or kinds of activity (ICT, technology, systems);
- Kinds of thinking (thinking skills and problem solving); and
- Personal skills, qualities, or attributes.

On one hand, all of the conceptions in the table in Appendix 1 (other than the OECD DeSeCo⁷ project) pick up the basic enabling skills, and also include reference to the more

⁷ Delineation and Selection of Competences

cognitive conceptions of research and problem-solving skills. All refer to technology, or more specifically to information technology, but the scholastic content areas are barely mentioned in any of them. Science is referred to in only the South African outcomes, and learning about culture and society was an afterthought as a (Mayer) key competency, 'Cultural Understanding'. Such notions are not found in the other conceptions.

Thus, it is fair to conclude that the notion of work-related skills bears little relationship to specific curriculum constructs or academic domains. In most cases they are concerned with skills rather than knowledge, and the skills are usually generic and transdomainal (or, at least, transdisciplinary—they go across the disciplines that make up the curriculum).

It is noticeable that all sorts of things are described as skills in work-related skills constructs, but particular emphasis is placed on practical skills such as 'Planning and organising' and 'Teamwork', on more personal skills such as 'Self-management' and 'Learning', and on values and attitudes such as loyalty and resilience rather than on the so-called academic or more purely cognitive skills related to knowing and understanding and doing this or that in the disciplines.

There was a great deal of 'curriculum mapping' in the 1990s to see if the key competencies were to be found in the curriculum of upper secondary education. An investigation of the extent to which employability skills (as well as the generic academic skills) were evident in secondary curriculum and practice was undertaken by ACACA (2003). It was generally concluded from this work that, although such practical and personal constructs as 'Planning and organising' and 'Teamwork' were implicit in secondary curricula, they could be made more explicit. The ACACA review also noted that greater clarity in definitions of employability skills is needed, that pedagogies and assessment strategies need to be developed for them, and that performance levels need to be described. Proposing assessment strategies and performance levels are outcomes of the current project.

The key competencies asked for a new emphasis on practical and personal skills in teaching and learning in secondary schools. As well, the Mayer Committee intended that all students exiting an educational or training program at or equivalent to years 10–12 should be given an assessment of their level of performance on the key competencies. The committee envisaged a 'Record of Performance on Key Competencies' as information for potential employers and for use in credit transfer between educational institutions. National sampling in key competencies assessment was envisaged for monitoring standards.

Doubts and concerns were expressed about the proposals for assessing the key competencies. It is a useful exercise to revise what the critics said. Their arguments, as identified and summarised in McCurry and Bryce (2007), were:

- Generic or transdomainal skills and hence the key competencies do not exist;
- Competence is knowledge-based and context-specific;
- Generic skills are unassessable;
- The key competencies are narrowly vocational and they undermine vocational education;
- The key competencies are new, foreign and unassessable constructs;
- Performance cannot be generalised because it is inconsistent and domain-specific;
- Competence and expertise are knowledge-dependent; and
- Teachers cannot make judgments about such constructs.

In more specific terms, the critics argued that the key competencies could not be successfully assessed for the following reasons (McCurry & Bryce, 2007):

- Students perform in significantly different ways in different subjects;

- Different teachers will have quite different perspectives of students; and
- It is not possible or meaningful to generalise about students across the curriculum.

As generic skills, the key competencies were seen as at odds with the structure and organisation of schools. The cross-curriculum⁸ nature of the key competencies could not be accommodated in conventional curricula, and some believed they would distract from and undermine the learning of discipline-based knowledge. It was contended by others that educational institutions should be concerned with knowledge and skills and that adaptation to the workplace should be a matter of on-the-job training. Overall it seemed to some that proposals associated with the key competencies were overly ambitious and impractical.

Between 1994 and 1996 the Commonwealth Government invested \$20 million in some 80 projects, mainly in schools, through a national program of trialling the key competencies. A review of the pilot phase in 1996 reported positively of the competencies themselves, but noted that little progress had been made in assessing them (Ryan, 1997). In part, the lack of success in implementing key competencies can be traced to three broad sources: to a perceived conflict with assessment approaches developed for the Statements and Profiles initiative, to the late collapse of that initiative after much effort was expended on it, and to external factors, including industrial action in some jurisdictions (see state contributions reported in Lokan, 1997).

An examination of different examples of generic, work-related assessment regimes further highlights the challenging nature of assessing and reporting on such skills. While there is widespread interest in generic and work-related skills, there has been very little system-wide assessment and reporting of such skills. There were few attempts to assess the SCANS competencies in the US (Oliver et al., 1997) or the Canadian Employability Skills (Conference Board of Canada, 2000). There was little effective assessment and reporting of Key Competencies (Clayton et al., 2003). The assessment of Key Skills in Britain seems to have been too narrow and too cumbersome (Hodgson & Spours, 2000; Powell, Smith, & Reakes, 2003; Turner, 2002). Little direct action has resulted from the extensive conceptual work undertaken in the OECD DeSeCo project, although the inclusion of the problem-solving domain in the 2003 PISA and the proposed PIAAC assessments may be attributed to the developmental work undertaken in the DeSeCo project.

Conclusion

The review of the five evaluative criteria—validity, reliability, objectivity, feasibility and usability—suggests that if these criteria as conventionally construed were rigorously applied to various methods of assessment and reporting, few if any of the methods would meet the criteria. However, the need for compromise has already been identified in relation to the assessment of school subjects generally. Given the nature of generic skills, the need for compromise in assessment and reporting might be even greater. Furthermore, assessing achievement in any domain of learning has several functions. A clear purpose of the current project is to recommend assessment methods that can lead to valid, reliable and fair reporting of student achievement. Assessment also signals what is important and can lead, either directly or indirectly, to the development of higher levels of generic employability skills receiving greater attention. This presents an additional dimension to the evaluation of possible approaches to assessing and reporting on employability skills.

⁸ Using ‘cross-curriculum’ as adjective; it does not require ‘curricular’.

Chapter 3 – Performance Levels

Chapter 3 attends to the production of one of the five outcomes required of this study by DEST; that is, ‘Identify performance levels for each employability skill that are appropriate for students completing Year 12 in Australia’ (DEST 2007: 5).

The existence of described *performance levels*, one set per employability skill, is arguably the linchpin in assessing and reporting student achievement in the eight employability skills, which are repeated below for easy reference during the discussion in this chapter.

1. Communication skills that contribute to productive and harmonious relations between employees and customers;
2. Teamwork skills that contribute to productive working relationships and outcomes;
3. Problem-solving skills that contribute to productive outcomes;
4. Initiative and enterprise skills that contribute to innovative outcomes;
5. Planning and organising skills that contribute to long-term and short-term strategic planning;
6. Self-management skills that contribute to employee satisfaction and growth;
7. Learning skills that contribute to ongoing improvement and expansion in employee and company operations and outcomes; and
8. Technology skills that contribute to effective execution of tasks.

This chapter takes all the facets of each of the employability skills from the ESF, examines them, and isolates those facets deemed to be assessable in the senior secondary school.

Using a generic statement as the basis for composition, descriptions of performance levels for each skill are composed. These descriptors are written in the language of student behaviour.

So that judgments about the quality of student work can be validly absorbed into overall decisions about the student’s level of performance on an employability skill, descriptions of facet standards are also composed for three skills. These descriptors are written in the language of evidence of achievement (from student work).

Note about completeness

This chapter contains the complete story on the specification of performance levels for all eight skills and of facet standards for three skills (Communication, Teamwork, and Problem solving). Thus it delivers on the required outcome—described performance levels, one set per skill. As the descriptions of three performance levels per skill emanate from the operational definition of that skill (a definition which is non-negotiable), confidence in their applicability is quite high. The descriptions of facet standards emanate from conjuring up a vision of activities that students might undertake at the behest of teachers (i.e. assessments designed to bring forth evidence of achievement of a skill (or skills)). Confidence in their applicability would not be high until teachers became involved in the process over an extended period.

The described performance levels and the described facet standards are presented in this report as indicative of a process that could be employed for all assessable facets of all employability skills.

The section on each of the employability skills is expanded to include descriptions of activities that give insights into the contexts in which the foundational facets can be developed and assessed.

Employability skills for assessment

In *Employability skills for the future*, ACCI and BCA defined each of the employability skills as having particular purposes and operationalised those definitions through facets. They foreshadowed the need to customise the skills to the diverse contexts in which they would be applied and developed by noting (ACCI & BCA, 2002: xii) (emphases added) that:

- The **mix and priority** of these facets **would vary** from job to job.
- The **list of elements is not exhaustive** but rather reflects the information provided by the specific enterprises interviewed for this study.
- The **list of elements is indicative** of the expectations of employers.
- The **level of sophistication** in the application of the element **will depend on the job level** and requirements.

In order to engender acceptance in the schooling sector of the proposals for assessing the employability skills and to customise them to the senior-school context, the facets listed for each of the employability skills were evaluated in terms of the ease with which they could be identified, developed and assessed in school subjects.

Furthermore, it was recognised that there would be few other opportunities for the group of young people who complete secondary schooling but do not proceed to post-school education and training to participate in employability skills development. These young people are likely to move into entry-level employment positions, and for this reason the facets selected for immediate attention are considered to be foundational.

The remaining facets of each employability skill could be developed for assessment but these facets (labelled ‘applied’ in Table 2) would probably require additional resources because their development and assessment might depend on creating simulated work situations. The ‘foundational facets’ on the other hand can be developed and assessed within existing school contexts. The sound establishment of performance levels in the foundational facets would provide the basis for the subsequent development of the applied facets in work situations.

The classification of facets for each employability skill as foundational or applied is shown in Table 2. An extra column, ‘Comments’, has entries for three skills—Communication, Teamwork, and Problem solving—because the identified performance levels for these are further elaborated later in this chapter. The comments also refer to ‘applied facets’, which will require a different assessment program from the ‘foundational facets’.

Table 2: Employability skills – foundational and applied facets

Employability Skill	Facets	Comment
<p>COMMUNICATION activities ... that develop students' abilities to use the dominant modes of communication to achieve clear purposes recognising the needs of the intended audience</p>	<p>Foundational facets: Listening and understanding Speaking clearly and directly Writing to the needs of the audience Reading independently Using numeracy effectively</p>	<p>These facets are developed and are assessable through most current senior secondary subjects.</p>
<p>... and that are a basis for productive and harmonious relations between employees and customers</p>	<p>Applied facets: Sharing information Persuading effectively Negotiating responsively Being assertive Empathising Establishing and using networks Understanding the needs of internal and external customers</p> <p>Speaking and writing in languages other than English</p>	<p>These facets are likely to be evident in many senior secondary subjects and in other school-based activities. These facets could be evaluated in simulated workplace activities.</p> <p>This facet is above entry-level standard.</p>
<p>TEAMWORK activities ... that develop students' knowledge of and facility in collaboration</p>	<p>Foundational facets: Working as an individual and as a member of a team Knowing how to define a role as part of a team Identifying the strengths of team members</p> <p>Applying teamwork to a range of situations (e.g. futures planning, crisis problem solving)</p>	<p>These facets are developed and are assessable through many current senior secondary subjects.</p> <p>This facet may represent higher than entry-level performance.</p>
<p>... and that are a basis for productive working relationships and outcomes</p>	<p>Applied facets: Working with people of different ages, gender, race, religion or political persuasion</p> <p>Coaching and mentoring skills including giving feedback</p>	<p>These facets may be developed by many students in co-curricular activities.</p> <p>This facet is above entry level.</p>
<p>PROBLEM-SOLVING activities ... that develop students' abilities to recognise problems, to apply strategies and to reflect on solutions</p>	<p>Foundational facets: Applying a range of strategies to problem solving Developing practical solutions Showing independence and initiative in identifying problems and solving them</p>	<p>These facets should be apparent in most senior secondary subjects and are vehicles for developing basic problem identification, strategy selection, monitoring and reflection processes.</p>

Employability Skill	Facets	Comment
... and that are a basis for achieving productive outcomes	<p>Applied facets:</p> <ul style="list-style-type: none"> Developing creative, innovative solutions Using mathematics including budgeting and financial management to solve problems Solving problems in teams Applying problem-solving strategies across a range of areas <p>Testing assumptions taking the context of data and circumstances into account</p> <p>Resolving customer concerns in relation to complex projects issues</p>	<p>These facets are applicable to some senior secondary subjects.</p> <p>This facet lies at the intersection of two skills, which is appropriate, but in the interests of parsimony, is excluded from the foundational category.</p> <p>This is an applied facet that would require simulation.</p>
INITIATIVE & ENTERPRISE activities ... that alert students to opportunities for novel approaches to situations and for acting autonomously to achieve goals	<p>Foundational facets:</p> <ul style="list-style-type: none"> Generating a range of options Translating ideas into action Being creative Identifying opportunities not obvious to others 	No comments hereafter in this column as per explanatory note in text
... and that contribute to innovative outcomes	<p>Applied facets:</p> <ul style="list-style-type: none"> Initiating innovative solutions Adapting to new situations Developing a strategic, creative, long-term vision 	
PLANNING & ORGANISING activities ... that develop students' knowledge and skill in managing time, information and resources	<p>Foundational facets:</p> <ul style="list-style-type: none"> Collecting, analysing and organising information Managing time and priorities – setting timelines, coordinating tasks for self and with others Being resourceful Taking initiative and making decisions 	
... and that contribute to long-term and short-term strategic planning	<p>Applied facets:</p> <ul style="list-style-type: none"> Establishing clear project goals and deliverables Adapting resource allocations to cope with contingencies Allocating people and other resources to tasks Planning the use of resources including time management Participating in continuous improvement and planning processes Developing a vision and a proactive plan to accompany it Predicting – weighing up risk, evaluating alternatives and applying evaluation criteria Understanding basic business systems and their relationships 	

Employability Skill	Facets	Comment
SELF-MANAGEMENT activities ... that encourage students to plan, monitor and evaluate goals, actions and achievements	Foundational facets: Evaluating and monitoring own performance Having a personal vision and goals	
... and that contributes to employee satisfaction and growth	Applied facets: Having knowledge and confidence in own ideas and visions Articulating own ideas and visions Taking responsibility	
LEARNING activities ... that lead to the recognition of the importance of learning and of opportunities to learn in order to increase personal capability	Foundational facets Managing own learning Having enthusiasm for ongoing learning Being open to new ideas and techniques Being prepared to invest time and effort in learning new skills	
...and that contributes to ongoing improvement and expansion in employee and company operations and outcomes	Applied facets: Being willing to learn in any setting – on and off the job Acknowledging the need to learn in order to accommodate change Contributing to the learning community at the workplace Using a range of mediums to learn – mentoring, peer support and networking, information technology (IT), courses Applying learning to ‘technical’ issues (e.g. learning about products) and ‘people’ issues (e.g. interpersonal and cultural aspects of work)	
TECHNOLOGY activities ... that develop students’ facility with technology	Foundational facets: Having a range of basic IT skills Using IT to organise data Being willing to learn new IT skills Having the occupational health and safety knowledge to apply technology	
...and that contribute to effective execution of tasks	Applied facets: Applying IT as a management tool Having the appropriate physical capacity	

Towards a set of described performance levels

The production of a set of described performance levels occurred in three steps, which are summarised in Table 3 before being explained in detail.

Table 3: What was involved in the production of a set of described performance levels

Step 1	The eight employability skills were shaped as assessment constructs	
	For each employability skill...	Operational definition analysed
		Foundational facets specified
Step 2	Global descriptors were devised for each employability skill Used for reporting grade for a particular skill	
	For each employability skill...	Possible student activities envisaged
		No. of categories of performance identified
		Student performance/behaviour described at three levels
		Labels given to performance levels 1, 2, 3 for reporting on employability skills
Step 3	Prototype descriptors were devised for each foundational facet of three skills Used to judge performance level for a particular skill	
	For each facet...	Operational definition analysed
		Evidence of achievement from student activities envisaged
		No. of discernible facet standards identified
		Student work/performance described at three standards
		Labels not required for reporting on facets; only required during the grading process that precedes reporting

The following explanation should be read in conjunction with Table 3.

The first step in the production of a set of described performance levels was to shape the employability skills as constructs for assessment in schools. This was accomplished by analysing the operational definitions of the skills and specifying facets to be targeted for assessment at the point of exit⁹ from school.

In order to make the assessment of employability skills manageable in the classroom context, each employability skill was customised to that context by selecting a subset of facets of the skill. Facets were selected if they were overtly identifiable as skills that students would be required to enact in their subjects, if they were necessary precursors to the development of more advanced facets, and if they would be required of entry-level employees. The third of these criteria arose because the target group of students are those who will undertake senior secondary education and seek to enter the workforce without undertaking further education or training. We anticipate that these students will need to achieve a grade of at least 'Basic' on each performance level on all employability skills in order to meet employers' minimal performance expectations. Thus, two categories of facets were distinguished, as shown in Figure 3.

The second step in the production of a set of described performance levels was to devise global descriptors for each skill in terms of what senior secondary students might be expected to achieve. This was accomplished by describing performance at three levels—basic,

⁹ See later for explanation of the 'exiting' student vis-à-vis the Year 12 student.

creditable, and advanced. These labels were chosen to emphasise the nature of the differences in achievement that need to be discerned by teacher–assessors.

The third step in the production of a set of described performance levels was to describe, for each foundational facet of each skill, the criteria by which evidence of achievement might be elicited and its quality judged. This was accomplished by envisaging activities that students might undertake and then describing performance on each facet at three standards. The words used to label these standards could be the same as for the performance levels (Advanced, Creditable, Basic), or simply just Standard 1, 2, and 3 from lowest to highest.

The entities for the assessment of employability skills are:

- Performance levels
- Performance-level descriptors
- Facet standards
- Facet-standard descriptors.

The entity for the reporting of employability skills achievement is:

- The performance level.

The two-stage process for arriving at a grade for reporting is discussed later, as is the way of displaying the results for a particular student.

Figure 3 summarises the relationships described above, building on the terminology introduced. It is provided as a referent for the text that follows it. The wording for the descriptions of performance levels and facet standards is proposed later in this chapter.

Foundational facets			
Applied facets			
Employability Skill	Levels	Labels for Levels	Description of Level
	Level 3	Advanced	Text that describes performance that minimally meets the highest level, (i.e. not the highest possible conceived performance) and clearly differentiated from Intermediate performance
	Level 2	Intermediate	Text that describes this level of performance in positive terms and that is clearly different from Basic performance
	Level 1	Basic	Descriptor for basic achievement that is a positive statement of what a student who operates at the minimally acceptable level can do.
Facet	Standards	No labels required Standards not reported	Description of Standard
	Standard 3		Text describing what an assessor might observe in the work of a student at the highest level of achievement on the facet
	Standard 2		Text describing what an assessor would observe in a student’s work when that student is demonstrating the facet at a level clearly higher than basic
	Standard 1		Text describing what an assessor might observe for a student who demonstrates the minimally acceptable achievement in relation to the facet being assessed

Figure 3: The identification of performance levels – three per skill, referenced to advanced, creditable and basic achievement

Illustration

The Employability Skill ‘Communication’ is now used to illustrate the concepts and terminology introduced into Table 2.

The *operational definition of the Employability Skill ‘Communication’* from the ESF is: Communication skills that contribute to productive and harmonious relations between employers and customers. The thirteen *facets of the skill* of Communication given in the ESF are listed below. The five facets targeted for assessment in schools are the so-called foundational facets.

Foundational facets

Listening and understanding
Speaking clearly and directly
Writing to the needs of the audience
Reading independently
Using numeracy effectively

Applied facets

Sharing information
Persuading effectively
Negotiating responsively
Empathising
Being assertive
Establishing and using networks
Understanding the needs of internal and external customers
Speaking and writing in languages other than English

The process of writing descriptions of standards included visualisation of activities that provided evidence of the skill being assessed. For Communication, the visualised activities develop students’ abilities to use the dominant modes of communication, recognise the needs of the intended audience, achieve clear purposes of the communication, and act as the basis for productive and harmonious relations between employees and customers.

Rationale for describing facet standards

The global descriptors at three overarching levels of performance for a given skill are, strictly speaking, meta-descriptors. Continuing with ‘Communication’ as the example, the three performance-level descriptors at Levels 1, 2 and 3 (3 being the highest level) reflect increasing competence in communicating; that is, in listening, speaking, writing, reading and so on as in the bold type in ‘Communication activities’ below, an extract from Table 2.

Communication activities
... that develop students’ abilities to **use the dominant modes of communication** to achieve clear purposes recognising the needs of the intended audience

... and that are a basis for productive and harmonious relations between employees and customers

If the level descriptors for overall performance (summative assessment) as they stand were to be applied to assessment instruments as they stand, the criteria for assessment of student work would be too broad—representing overarching properties like ‘(student) independence and degree of (environmental) structure [in communicating in multiple modes]. Such broad and overarching criteria need to be customised to the each skill to enable the specific skills to be assessed in student work. Without this customising, the validity of the assessment would be compromised.

Performance levels and evidence

It is envisaged that the report on a student’s employability skills achievement would be in the form of one result (grade) per skill. This would be the case whatever the assessment instruments used, whatever the reporting format used, and whatever authority is responsible for the reporting (all of which are canvassed in later chapters in this report). The grade could be expressed (coded) in various ways such as letters, numbers, words, and/or descriptors.

Table 4: Options for reporting performance levels as grades

Performance level	Options (Choose at least one.)				
	Use words	Use numbers	Use other words (list not exhaustive)	Use letters	Use descriptors alone, or as a legend to numbers or letters
Level 3	Advanced	3	High, distinguished, ...	A	Student independently completes complex tasks in multiple environments.
Level 2	Creditable	2	Intermediate, commendable, ...	B	Student completes tasks in mainly unstructured environments with minimal teacher help and support.
Level 1	Basic	1	Acceptable, sound, ordinary, ...	C	Student completes routine tasks in structured environments usually with teacher help and support.

The reader should return to this table when noting Figure 5, later, which displays a possible reporting format.

Two-stage process

Before it is possible to communicate results on a particular employability skill to a variety of audiences (i.e. to report), it is necessary to assess the student’s overall level of performance on that employability skill. Arriving at an overall (or exit) assessment is a two-stage process.

In the explanation that follows, the volume of words belies the simplicity of the process.

What is important about the process is **that** validity and reliability of the assessment requires two things:

1. Student work that emanates from good assessment instruments; and
2. Teacher–assessors who have a ‘feel for’ the quality implicit in the descriptions of performance levels and facet standards.

The two stages for assessment of a particular employability skill are:

Stage 1: Judging student work against facet standards

Look at student work/performance (i.e. evidence of achievement in the employability skill). For each of the (foundational) facets, look at the three described standards. Judge which of the descriptions best matches the student work/performance.

Stage 2: Mapping facet standards to performance level

Decide on the overall grade for reporting (Performance Level 1, 2, or 3) by combining the assigned Facet Standards from Stage 1. Methods for doing this are described after Figure 5.

The process is now exemplified. The teacher–assessor uses the matrix in Figure 4 for assessing the student Bertha Mason on the employability skill, Communication.

A cross (×) in a cell on each row of the matrix represents the outcome of Stage 1 above.

The entries in the last two rows of the matrix represent the outcome of Stage 2 above. Methods for deriving this outcome are described after Figure 4.

Skill: Communication	Student name: <i>Bertha Mason</i>		
Facet	Facet Standard		
	1	2	3
Listening and understanding	Listen to and appropriately respond to speakers' views during discussion x	Listen to and recognise the ways in which speakers express or imply a point of view and values	Understand and illustrate the relationship between purpose, form, language and audience in a range of written and multi-modal text types
Speaking clearly and directly	Understand the ways oral language can be structured for some audiences and purposes Use tone and vocabulary appropriate to the intended purpose of a (given) text	Understand the ways oral language can be varied for different audiences and purposes Use tone and vocabulary appropriate for a range of different text types x	Understand the relationship between purpose, form, language and audience in a broad range of oral expression Effectively use tone and vocabulary appropriate for a range of different text types
Writing to the needs of the audience	Understand and show the ways written language can be structured for some audiences and some purposes	Understand and show the ways written language can be varied for different audiences and different purposes x	Understand and illustrate the relationship between purpose, form, language and audience in a range of written and multi-modal texts
Reading independently	Extract literal meaning in everyday or standard written text	Extract implied meaning in most text modes	Recognise the value of the structures, features and conventions used by authors for constructing meaning in a range of multi-modal texts x
Performance level overall: 2			
Grade for reporting: <i>Creditable achievement</i>			

Figure 4: Two-stage process for arriving at overall grade for a particular skill

Analytical method for combining results

Use a simple algorithm to decide the overall performance level.

Two possibilities:

(i) Facets are equally weighted.

Overall performance level is the average of the facet standards.

Student in Figure 4 would then be at Level 2 in Communication:

$$\text{Average of facet standards} = (1+2+2+3)/4 = 2$$

(ii) Facets are unequally weighted.

Overall performance level is the weighted average of the facet standards.

Appropriate when one facet is deemed to be more important than another.

Say, in the overall assessment of Communication, the facet 'Reading independently' is deemed to be worth three times as much as any other facet.

$$\text{Weighted average of facet standards} = (1 \times 1 + 1 \times 2 + 1 \times 2 + 3 \times 3)/(1+1+1+3)$$

$$= (1+2+2+9)/6$$

$$= 14/6$$

$$= 2.3$$

Student in Figure 4 would still be at Level 2 in Communication.

A software tool, perhaps based on a spreadsheet, could be developed to automate the process of calculating scores and generating grades. The tool could incorporate rules for rounding off.

Holistic method for combining results

Make an on-balance judgment to decide the overall performance level.

Two possibilities:

(i) Facets are equally valued.

Based on an impression of the pattern of the facet standards, decide on the overall performance level, ensuring that a high standard on one facet compensates for a lower standard on another.

(ii) Facets are unequally valued.

If one facet is deemed to be more important than another, proceed as above, but ensure that standards on the three facets contribute to the overall performance level in a manner reflective of their hierarchical positions. For example, if the student is Standard 1 on a facet deemed to be unimportant relative to the others, do not allow this to have much effect on the overall grade.

Audit of processes used for setting, describing and representing performance levels

This section checks the application of design 'rules' to the identification and elaboration of performance levels for employability skills. The left-hand column provides the short form for rules gleaned (Matters, 1999) from the measurement literature and best practice on setting, describing and representing performance levels (or achievement standards). The right-hand column provides a check on the application of those rules to the identification of performance levels for employability skills as undertaken in this study.

Table 5: Application of design rules to the production of described levels of performance

The design rule	Applied?
Setting levels of performance	
The levels reflect the legitimacy of the curriculum experience of students at school in Australia.	Yes
The levels represent the demands of the employability skills although not necessarily encompassing the totality of the skills (all possible facets).	Yes
The levels are realistic and attainable by the range of senior secondary students in Australia.	Yes
Three levels per facet represent the expected categories of student performance sufficient to enable differentiation of performance.	Yes
Describing performance levels	
The level descriptors are fresh statements and not mere replications of employability skill or facet definitions.	Yes
The descriptor for the highest level gives students something to aspire to.	Yes
The descriptor for the lowest level is not written in deficit terms.	Yes
The descriptors are written in the language of the employability skills and their assessable facets.	Yes
Three performance-level descriptors cover a range of performances on a particular employability skill.	To be tested out
Three facet-level descriptors cover a range of performance on a particular facet.	To be tested out
The level descriptors clearly describe the qualities of each of the performances in the range.	Attempted to do so
The level descriptors are of an appropriate grain size. That is, they are not so coarse that the categories are indistinguishable; and they not so fine that the required precision of the assessment decisions is impossible based on the evidence available.	This was the intention.
Representing performance levels	
Are the written level descriptors able to be found in real student work?	Not yet known
Are the level descriptors categorised into constructs that justify the range and balance of skills experienced in learning?	Assumed so
Have the level descriptors as set and described been tested out with student work and revisions made on the basis of data gathered in the exercise?	Not yet

Establishing levels is conditional

Although a hierarchy of levels has been identified at a specified juncture (exit from Year 12) and levels of performance have been identified and descriptors of intended levels composed, these levels require verification in practice before they can be said to be established, because the primary evidence of learning (i.e. student work) that purports to be at a described level has not yet been seen.

Although the assessments are essentially criteria-based, there is a flavour of ‘norming’; that is, the levels have a sense of ‘absoluteness’ because, in reality, they could be applied to anybody from kindergarten to old age. The entry-level descriptor is pitched at Year 10. The expected set of competencies is for students who have reached Year 10 even though the reporting juncture was to be in (or perhaps at the end of) Year 12 (also refer back to discussion of targeted students in Chapter 1).

The expertise of teacher–assessors in reaching agreement as to whether the evidence does indeed match the description of the performance level has not yet been demonstrated.

The process of devising performance levels has not involved external experts in employability skills development and assessment. Teachers who are able to envision student work should be involved in identifying performance levels over a period of months if assessment and reporting against national levels is to be implemented.

Form of representation chosen

Performance-level descriptors are capable of being represented in a variety of formats such as a chart, matrix, dimension, or pole. The simple matrix is chosen here. It is suitable for the representation of three levels as discrete categories of performance.

Note on status of descriptors

This report fulfils the project requirement to describe performance levels for each of the eight employability skills. In addition, it provides standards for the foundational facets of three employability skills (‘Communication’, ‘Teamwork’, and ‘Problem solving’). The facet standards for these three skills are indicative of what will be required for the foundational facets of the remaining five employability skills in order to achieve robust assessment and reporting of them.

The descriptions of three performance levels per skill emanate from the operational definition of each skill (specified in the ESF). Confidence in their applicability is high. The elaborations of performance levels into facet standards, however, emanates from envisaging activities that students might undertake at the behest of teachers (i.e. assessments designed to bring forth evidence of achievement of each skill). Confidence in their applicability cannot be high until teachers become involved in the process over an extended period. That is, validation of the facet standards is required in practice.

It should also be noted that the sections on each of the employability skills have been expanded with descriptions of activities that give insights into the contexts in which the foundational facets can be developed and assessed.

Evaluation of descriptors

A design rule in setting, describing and representing levels or standards is that differentiation between standards involves two variables, element and degree, applied separately or together. The level descriptors composed for this report meet this requirement. Again, the employability skill, ‘Communication’, is used as the example.

The element is ‘completing a communication event’; there are several degrees of completing and several shades of event: the degree to which the situation is novel (cf. routine); the degree to which the event is structured (cf. unstructured); and the level of support (low, medium,

high). These give clues to qualifiers to distinguish levels or standards. The use of adjectives and adverbs makes it easier for the writing process to have consistent syntax.

Description of performance levels for an employability skill

For a given employability skill there are three performance levels (for grading and reporting) plus an ungraded category.

Employability Skill: Communication

The foundation of communication is the ability to use the dominant modes of communication to achieve clear purposes while recognising the needs of the intended audience. Good communication is the basis for productive and harmonious relations between employees and customers.

Communication involves:

- Writing to the needs of the audience
- Speaking clearly and directly
- Listening and understanding
- Reading independently
- Using numeracy effectively.

Performance level	Grade	Description
3	Advanced achievement	Student independently completes complex tasks in multiple environments.
2	Creditable achievement	Student completes tasks in mainly unstructured environments with minimal teacher help and support.
1	Basic achievement	Student completes routine tasks in structured environments usually with teacher help and support.

Employability Skill: Teamwork

The foundation of teamwork is understanding and having facility in collaboration. Good teamwork is the basis for productive working relationships and outcomes.

Teamwork involves:

- Being aware of the different views and perspectives of others
- Being able to adapt to group processes
- Applying teamwork skills to a range of situations.

Performance level	Grade	Description
3	Advanced achievement	Student works cooperatively and harmoniously, while demonstrating tolerance and leadership as needed in pursuit of the expected outcomes of an activity and within time constraints.
2	Creditable achievement	Student works cooperatively and harmoniously with others on activities with expected outcomes.
1	Basic achievement	Student works cooperatively and harmoniously in a group on a prescribed activity with a modest level of demand.

Employability Skill: Problem solving

The foundation of problem solving is the ability to recognise problems, apply strategies, and reflect on solutions. Good problem solving is the basis for achieving productive outcomes.

Problem solving involves:

- Accurate identification of the problem
- Taking a practical approach to solving problems
- Reasoning logically and critically in pursuit of solving the problem.

Performance level	Grade	Description
3	Advanced achievement	Student demonstrates an extensive understanding of problem complexity and solution strategies. Consistently selects an appropriate strategy to arrive at the <u>best</u> solution.
2	Creditable achievement	Student demonstrates general understanding of the nature of problems and problem identification. Can usually select an appropriate strategy to arrive at <u>a</u> solution.
1	Basic achievement	Student demonstrates basic understanding of a simple (given) problem, applies practical or routine strategies, and arrives at <u>a</u> solution.

The other five Employability Skills

Initiative & Enterprise

The foundation for initiative and enterprise is the recognition of opportunities for novel approaches to situations and for acting autonomously to achieve goals. Possessing initiative and enterprise is the basis for achieving innovative outcomes.

Initiative and enterprise involve:

- Taking a positive approach to challenges and opportunities
- Being flexible and adaptable
- Identifying options and possibilities.

Performance level	Grade	Description
3	Advanced achievement	Student seeks challenges and opportunities, judges when it is appropriate to be flexible and adaptable, and creates options and possibilities.
2	Creditable achievement	Student recognises opportunities, takes a positive approach to challenges, is prepared to be flexible and to adapt to challenges and opportunities, and identifies a range of options and possibilities.
1	Basic achievement	Student responds positively to some challenges, adapts to challenges in some circumstances, and identifies some options and possibilities.

Planning & Organising

The foundation for planning and organising is knowledge about and skill in managing time, information and resources. Good planning and organisation is the basis for contributing to long-term and short-term strategic planning.

Planning and organising involve:

- Managing one's own time and priorities
- Organising information
- Planning projects/tasks.

Performance level	Grade	Description
3	Advanced achievement	<p>Student analyses and responds to time-management challenges by drawing on experience, takes responsibility for managing own time, and gives guidance to others.</p> <p>Imaginatively analyses and organises information by drawing on experience of organising in other contexts, in ways that might enable novel but appropriate practical solutions.</p> <p>Envisages multiple ways of accomplishing a project/task, draws on experience to select the best pathway, and plans to deal with unforeseen risks and contingencies.</p>
2	Creditable achievement	<p>Student organises own time so as to set priorities and goals, predicts outcomes and adjusts plans.</p> <p>Analyses and organises information in an appropriately logical and practical arrangement.</p> <p>Foresees and plans for some risks and sets logical steps to achieve goals or meet deadlines, including planning the use of resources.</p>
1	Basic achievement	<p>Student recognises the need to organise time and understands the consequences of not doing so.</p> <p>Organises straightforward information (such as a series of tasks), mainly following and applying learned guidelines.</p> <p>Breaks down a task into steps to meet deadlines with guidance.</p>

Self-management

The foundation for self-management is the ability to plan, monitor and evaluate goals, actions and achievements. Good self-management is the basis for contributing to employee (self) satisfaction and growth.

Self-management involves:

- Setting goals and taking responsibility for oneself
- Evaluating and monitoring one's own performance
- Understanding one's own strengths and weaknesses.

Performance level	Grade	Description
3	Advanced achievement	Student sets challenging goals, takes responsibility for self in difficult situations, effectively evaluates and monitors own performance and builds on feedback, and makes the most of strengths and minimises weaknesses.
2	Creditable achievement	Student sets realistic goals, takes responsibility for improving own performance, evaluates and monitors own performance, and understands own strengths and weaknesses.
1	Basic achievement	Student makes plans, sets short-term goals, understands and interprets feedback about own performance, and identifies own strengths and weaknesses.

Learning

The foundation for learning is the recognition of the importance of learning and opportunities to learn in order to increase personal capability. Being good at learning is the basis for ongoing improvement and expansion in employee and company operations and outcomes.

Learning involves:

- Taking a positive approach to learning
- Reflecting on one’s own learning and identifying learning choices
- Applying learnings in new contexts.

Performance level	Grade	Description
3	Advanced achievement	Student seeks and makes opportunities for learning, reflects realistically and effectively on own learning, and recognises and understands the need to adapt learnings for application in different contexts.
2	Creditable achievement	Student takes a positive approach to learning, reflects on own learning and identifies learning choices, and applies learnings in new contexts.
1	Basic achievement	Student takes opportunity to learn, recognises and uses learning opportunities, and learns new things in familiar contexts.

Technology

The foundation for this skill is facility with the technology. Being good with technology is the basis for the effective execution of tasks.

Using technology involves:

- Possessing a range of IT skills
- Transferring understanding from one technology or application to another
- Having a positive approach to technology, including using it responsibly, with knowledge of Occupational Health & Safety.

Performance level	Grade	Description
3	Advanced achievement	Student uses a broad range of IT skills, understands the principles that allow transfer of understanding from one technology or application to another, seeks technological experiences and challenges, and attempts to deal with OH&S issues.
2	Creditable achievement	Student uses a range of IT skills, transfers understanding from one technology or application to another, takes a positive approach to technology, and understands OH&S issues.
1	Basic achievement	Student uses some basic IT skills, understands the underlying principles of ICT, takes a positive approach to some technology, and recognises that problems can arise from technology.

Descriptions of facet standards for an employability skill

Employability skill: Communication

Facet: Writing to the needs of the audience

Writing to the needs of the audience involves:

- Understanding the ways writing can be structured for different audiences and purposes
- Using vocabulary and sentence structures appropriate to the intended purpose of the text
- Using strategies for drafting and redrafting for audience appropriateness, prioritising and sequencing ideas
- Proofreading and editing writing for accuracy, consistency and clarity.

Facet standard	Description
3	Understands and illustrates the relationship between purpose, form, language and audience in a range of written and multimodal text types.
2	Understands and shows the ways written language can be varied for different audiences and purposes.
1	Understands and shows the ways written language can be structured for some audiences and purposes.

Facet: Speaking clearly and directly

Speaking clearly and directly involves:

- Understanding the ways oral language can be structured for different audiences purposes
- Using oral language appropriately and expressively.

Facet standard	Description
3	Understands the relationship between purpose, form, language and audience in a broad range of oral expression. Effectively uses tone and vocabulary appropriate for a range of different text types.
2	Understands the ways oral language can be varied for different audiences and purposes. Use tone and vocabulary appropriate for a range of different text types
1	Understands the ways oral language can be structured for some audiences and purposes. Uses tone and vocabulary appropriate to the intended purpose of a (given) text.

Facet: Listening and understanding

Listening and understanding involves:

- Listening and responding in oral exchanges
- Analysing relationships between texts, contexts, speakers and listeners
- Interacting, verbally and non-verbally, with speaker to the extent of showing understanding or seeking clarification.

Facet standard	Description
3	Evaluates oral presentations and the ways in which speakers use vocabulary and tone of voice to influence/persuade/argue etc.
2	Listens to and recognises the ways in which speakers express or imply a point of view and values.
1	Listens to and appropriately responds to speakers' views during discussion.

Facet: Reading independently

Reading independently involves:

- Comprehending written language
- Making inferences (i.e. detecting implications and nuances)
- Appreciating the way different texts are structured for different audiences and purposes.

Facet standard	Description
3	Recognises the value of the structures, features and conventions used by authors for constructing meaning in a range of multimodal texts.
2	Extracts implied meaning in most text modes.
1	Extracts literal meaning in everyday or standard written text.

Representation of profile of results

Figure 5 represents the eight employability skills in a non-hierarchical way. A case could of course be made that there is a hierarchy and therefore a simple list is appropriate; for example, always having ‘Communication’ at the top of the list as the employability skill of all skills.

Another way of representing the skills in a non-hierarchical way is to simply list them in alphabetical order or in the order in the original ESF (as per Table 2 earlier in this chapter).

Teachers and students should be able to bring to mind the complete list without much trouble just as teacher–assessors should be able to bring to mind all the criteria being used for assessment without much trouble. Unless there is a mnemonic¹⁰ for memorising the list of employability skills¹¹, a diagram might be effective. Figure 5 is one possible representation of the eight employability skills in the ESF.

Legend

- C Communication
- TY Technology
- PS Problem Solving
- IE Initiative & Enterprise
- PO Planning & Organising
- S Self-management
- L Learning
- TW Teamwork

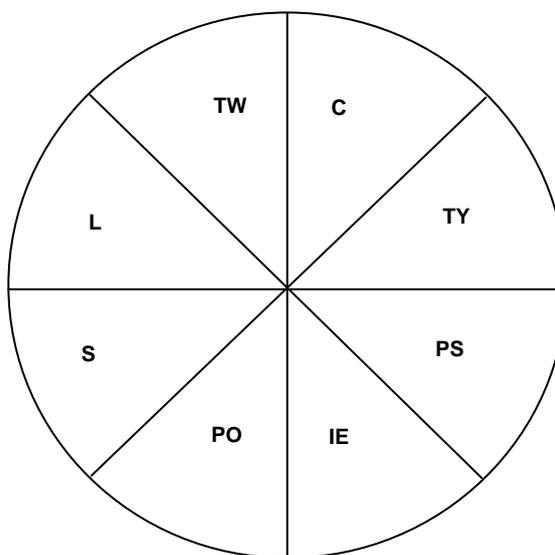


Figure 5: Diagrammatic non-hierarchical representation of employability skills

¹⁰ As in ‘Every Good Boy Deserves Fruit’ in order to recall the lines of the treble clef (EGBDF)],

¹¹ Many versions could be composed for C, T, P, I, P, S, L, T (which is the order of the employability skills in the ESF. Here is one: Cats Tell People If Pigs See Light Truly. This mnemonic does not differentiate the 2 T’s (Technology and Teamwork) and the 2 P’s (Problem solving and Planning & Organising). Going to the second letter would work for the 2 P’s but not for the 2 T’s. And so on.

Reporting

Figure 6 presents a display of the results of employability skills assessment that gives an immediate sense of the different patterns or constellations of results from student to student. An overlay of Figure 5, it is a special type of bar graph (after Florence Nightingale) where the bars are segments of a circle.

The template for the circle would have the radii estimated such that the ratio of the areas of the large, medium and small 'slices' is 3:2:1. The design specifications for the template would need to specify the use of colours. There could be three colours, one for each of the performance levels. There could be eight colours, one for each of the employability skills as in Figure 6. Where there is no colour in a slice or segment (i.e. that eighth of the circle/wheel is blank), the student is ungraded for that particular skill.

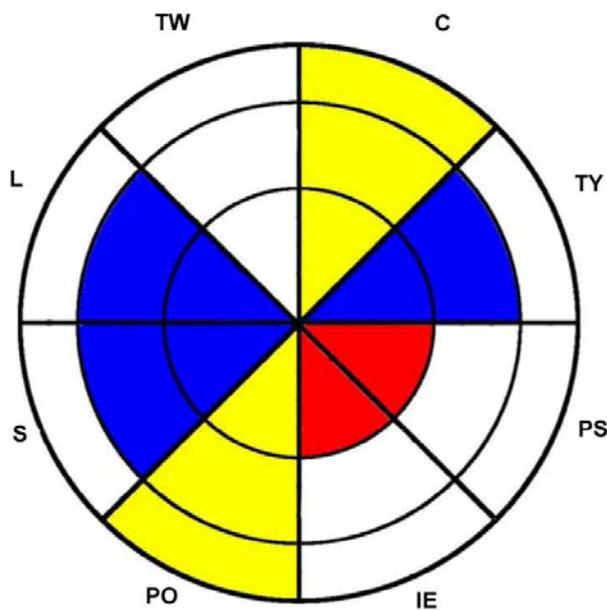


Figure 6: Diagrammatic representation of results for reporting

Interpretation of the diagram for a particular student:

The diagram shows a *profile of results* for employability skills achievement. This student has been awarded the following grades for her/his achievement in the employability skills:

Employability skill	Result
Communication	Advanced
Technology	Creditable
Problem solving	Basic
Initiative & Enterprise	Basic
Planning & Organisation	Advanced
Self-management	Creditable
Learning	Creditable
Teamwork	Ungraded

Conclusion

Two processes can be designed. There is a process for identifying and describing a set of performance levels for use by teacher–assessors. And there is a process for assessing and reporting on student achievement using a set of performance-level descriptors supported by a set of facet-standard descriptors for each employability skill.

The starting point for working towards a set of described performance levels is the assumption that it is possible to discern qualitative and quantitative differences in student work such that student performance on a given employability skill falls into one of three categories, Level 1, Level 2, or Level 3 (highest). Although the levels are only tentatively established, the codes for reporting student performance at Levels 1, 2 and 3 are Advanced, Creditable, and Basic, respectively. The code ‘Ungraded’ is used when, for whatever reason, no evidence of achievement is forthcoming.

Assigning a performance level to a student for reporting purposes could be effected through an evidence-based two-stage decision-making process supported in the first stage by facet-standard descriptors and in the second stage by performance-level descriptors.

Reports on individual student achievement in the employability skills could be in the form of coloured diagrams that highlight the variability in skills within and across students.

Chapter 4 – Models and Methods for Assessing and Reporting Employability Skills Achievement

Models for assessing and reporting employability skills have been found through a literature and web search of generic and employability skills initiatives in secondary schools, vocational education and training providers, and higher education institutions in Australia and overseas. Selected examples are presented in Appendix 3.

The examples can be classified into the following main groups:

- Standardised tests;
- Common assessment tasks;
- Teacher-generated tasks (including performance assessment);
- Judgment by teacher groups;
- Embedded development/assessment;
- Portfolio construction; and
- Self-assessment.

Reasons for the separation into the main categories arise from the development of a framework for model specification, discussed below.

Descriptions of assessment methods

Standardised tests

Standardised tests comprise items for which students select responses from prescribed options (typically multiple-choice items) or for which students provide limited constructed responses. These items are developed according to item specifications that include the particular constructs to be tested and the scope and range of abilities that are being assessed. The broad scope of a test is subdivided into small units of knowledge or skill and each element is assessed by a number of discrete items. Test items are evaluated by expert panels then trialled and are accepted, rejected or modified before final testing.

Where constructed response items are included, for example items eliciting short answers, raters are trained to recognise key features in responses and multiple raters are used, at least on a subset of scripts, to check that inter-rater reliability reaches an acceptable standard.

Although many judgments are made in delimiting the scope of the test and in prescribing the scope and range of items, variability in grading responses is either eliminated (for multiple-choice items) or minimised (in grading constructed responses).

The results of trial and final tests are analysed statistically to ensure that acceptable criteria are achieved, usually in relation to reliability, but validation studies are also conducted to check that test scores correlate with constructs that are thought to be related and to other criterion performance measures.

Example: The Programme for International Student Assessment (PISA) project is an example of the processes that surround standardised testing. Sample test items are available and reports based on the testing are published (see, for example, Thomson & de Bortoli, 2007).

Common assessment tasks

Assessment tasks are developed that provide an opportunity for students to demonstrate a range of related abilities that constitute a complex cognitive ability. Responses to these tasks have multiple dimensions, and judgments are made about each dimension according to performance-level descriptors.

Example: Rich Task, Year 9, Opinion-making Oracy
(<http://education.qld.gov.au/corporate/newbasics/pdfs/yr9rt6.pdf>)

Teacher-generated tasks

These are tasks that are selected or developed by individual teachers. They may be related to a particular subject and indeed could be an existing assessment activity in a subject. In the employability skills assessment context, these would be tasks that enable students to develop and demonstrate facets of an employability skill. Ideally, such tasks would be similar to the common assessment tasks described above. In practice, it is likely that few teachers would have the time to develop tasks with the degree of supporting material and the assessment guidelines that are available with centrally-developed tasks.

Example: Matters (2005).

Judgment by groups of teachers

Teachers meet and consider the employability skills of individual students whom they have taught or otherwise interacted with in co-curricular activities during a school year. Teachers consider each employability skill in turn, and describe the evidence they have been able to gather that illustrates each student's achievement of that skill. This process may be supported by a rubric describing the behaviours likely to be observed in students performing at different levels of the targeted employability skill.

Example: McCurry and Bryce (2000).

Embedded development and assessment

Embedded development and assessment of employability skills occurs when elements of existing subjects are mapped onto facets of the ESF. Typically, the existing assessment activities that are selected demand sufficient facets of an employability skill to enable a judgment about that employability skill to be made while achieving the original purpose of the assignment. Typically, also, judgments are scaffolded through the use of performance-level descriptors or other rubrics.

Example: Curtis and Denton (2003).

Portfolio construction

Portfolio construction is distinguished from portfolio assessment. The construction of a portfolio is the selection and aggregation by individuals of evidence of their own achievement of particular skills (which may include employability skills). Portfolios may be paper-based or electronic. Electronic versions are popular because of the ease with which they can be updated. Templates can be provided, and electronic portfolio systems may have facilities, such as filters, to enable students to select and present views of the contents in print or electronic form for specific purposes and audiences.

Two approaches to portfolio assessment are described by Troper and Smith (1997). They expressed some reservations about the validity and reliability of portfolio assessments. One of the problems is that if the portfolio is assessed as an object along with its content, then a construct other than the target one is brought into the judgment, threatening the validity of the assessment. This problem is compounded if different templates provide different levels of support to students. Assessing portfolios is time-consuming, and if they are to be useful resources to students, they would need to be assessed repeatedly with feedback provided to guide improvement.

Example (of a template): Conference Board of Canada (2000).

Self-assessment

Self-assessment is more often advocated than instigated (Biggs & Moore, 1993; Boud, 1995, 2002; Sadler, 1989; Wiggins, 1998). It is, however, an element of the ability-based curriculum of Alverno College, for example (Loacker, 2000). The advocacy of self-assessment is based on the argument that a capacity for monitoring and reflecting on their own work is necessary for students to be able to improve their performance. Sadler (1989: 121) expresses this view:

Stated explicitly, therefore, the learner has to (a) possess a concept of the standard or goal, or reference level being aimed for, (b) compare the actual (or current) level of performance with the standard, and (c) engage in appropriate action which leads to some closure of the gap.

Thus, for self-assessment to be viable for students' learning and improvement, performance standards must be expressed in terms that students can understand. However, this is not an adequate basis for this form of assessment to lead to reports of achievement.

Example: Loacker (2000).

An employability skills assessment model

An analysis of the cases for each category of assessment types led to the development of several criteria against which each type might be assessed. These criteria are elaborated below as a series of questions.

- What purpose is served by the form of assessment?
- Who controls the setting, marking and certification of the assessment?
- What degree of standardisation is inherent in the assessment?
- To what extent is the assessment of employability skills implicit or explicit?
(A related issue is the extent to which teaching/delivery of employability skills is embedded in, or separate from, existing curriculum and instructional practices.)

These characteristics, which are represented variously as polarities, stages and continua, are presented in Figure 7. The characteristics are discussed briefly then a framework for identifying models of assessment is developed.

Purpose			
Assist learning	Assess individual achievement	Program/system evaluation	
Control of Assessment			
Individual teacher	School	State	National
Degree of Standardisation			
Unique		Standardised	
Specificity of Assessment			
Inferred		Explicit	

Figure 7: Characteristics of assessment

Fitness for purpose

In Chapter 2 (Fundamental Issues and Concepts), three broad purposes were recognised for assessment, namely to measure student achievement, to plan and assist learning and to evaluate programs. In that discussion, attention was drawn to backwash effects that may arise incidentally from chosen assessment approaches. Assessing a domain signals its importance

and this may lead to greater attention being paid to teaching related content. Assessment may also lead to alternative ways of conceptualising and teaching existing subject content.

The main purpose of the assessment of employability skills is to generate a report of individual achievement of these constructs. However, merely reporting achievement will not directly lead to enhanced performance across the cohort of school leavers. Assessment that informs students and teachers of their current status against a graduated set of standards-based performance levels also informs them of the requirements of the next performance level. This information enables teachers to plan instruction that will lead to higher levels of student achievement. Thus, in evaluating assessment methods, it is necessary to consider their suitability both for measuring achievement and for assisting learning—their fitness for purpose.

Control of assessment

The types of assessment that are used, who marks the assignments, and who certifies the results provide a set of related control dimensions over assessment. Assessment control may be localised so that individual teachers decide what tasks will be the basis of an employability skills assessment, what standards will apply in judging students' performances, and therefore what grades are assigned to students on those tasks. Under these conditions, it is the teacher whose authority underpins the grade awarded. Localised teacher judgment of performance assessment or holistic judgment by individual teachers about students' achievement of employability skills based on observations of students in a variety of classroom activities is of this type.

Assessment activities and standards may be suggested and prescribed centrally, but the judgment of individual student performance may be left to individual teachers. An example of this approach is the curriculum resource materials and assessment standards and guidelines prepared and distributed in Scotland for the teaching and assessment of Core Skills. This assessment is certified locally, although the education system provides some warrant for the results reported.

Assessment activities, standards and the marking of student scripts may be centrally controlled. This is the case with the Graduate Skills Assessment (see Appendix 3). The test is secure, although sample items are made available; the conditions for taking the test are prescribed; and the scripts are collected and returned to a central agency for marking. The results of tests such as this have the authority of the central testing agency.

Degree of standardisation

The standardisation of tests and test items is related to the degree of central control or local autonomy over assessment standards, items and grading practices. Teachers may be free to design their own assessment activities, to adopt recommended activities with or without grading guidelines, or be required to use or select from a range of centrally designed test items and scoring standards.

Individual teachers may make judgments about students' demonstration of employability skills based on their observations of the students over a range of classroom-based and co-curricular activities. In this case, the activities that are the subject of the judgment are unlikely to be common for all students. Teachers' judgments integrate a diversity of observations and are quite non-standard.

In standardised tests such as PISA, considerable effort is expended during the development of test specification and test items to ensure that a common pool of items is available for use across all candidates. Items are assigned to students in ways that ensure students are treated fairly. What constitutes acceptable responses to items is decided in advance of test administration and all students' responses are scored against a common set of rules.

Following test administration, test data are analysed to ensure that items behave as intended and that item difficulty is taken into account in scoring students. All aspects of such tests are highly standardised.

It is worth noting that all forms of assessment, from the non-standardised to the highly standardised, involve professional judgment. In the case of holistic teacher judgment, the judgment occurs in the act of assessing students while in the highly standardised testing approaches, judgments are made during test, item and response specification.

Specificity of assessment

Assessment of employability skills may be inferred or explicit. Inferred assessment arises when particular employability skills are asserted to be implicit in existing curriculum materials. In the past, key competencies were mapped onto curriculum documents. It was asserted that if a student completed an element of the prescribed syllabus, and that element required a key competency, the student must have demonstrated that key competency and could be automatically credited with it. This was one of the ways in which key competencies were included in training packages and were recorded as being achieved.

Curriculum mapping need not be overt or systemic. Judgments about the achievement of employability skills can be left to individual teachers who can choose to equate aspects of employability skills with the content and processes of their classroom practices and could assert that, by doing these activities, students have demonstrated those aspects of the employability skills.

The assessment of employability skills can be quite explicit. Each employability skill can be described in terms of its knowledge and skill components. Rubrics can be constructed that identify each of the identified components and students can undertake activities in which those elements are sought and observed explicitly. An example of this approach is given in Curtis and Denton (2003). In this case, a generalised assessment tool was developed that identified a set of major problem-solving processes and indicators of these processes. Teachers were required to use evidence in students' responses to assessment activities to decide whether the component processes had been enacted and to score the problem solving performance explicitly but in parallel with the content assessment based on the same piece of work. This case also demonstrates the use of embedded, rather than separate, delivery of employability skills.

A framework for classifying assessment methods

Of the four characteristics of assessment approaches identified above—fitness for purpose, locus of control, degree of standardisation, and specificity of assessment—the two key attributes leading to a differentiation of assessment approaches appear to be degree of standardisation and locus of control. Figure 8 shows how the two key attributes are used to define a field within which models of assessment can be described.

The two dimensions of the model are closely related in that an assessment that is highly decentralised is likely also to be low on the standardisation dimension. Conversely, assessment practices that are highly standardised are likely to be strongly centralised.

An important consideration is the authority with which results of the assessment of employability skills are certified. For highly centralised and standardised assessment, it would be possible for employability skills achievement to be recognised at a state or national level and for the result to appear on a senior secondary certificate. In decentralised assessment models with a low degree of standardisation, the decision may be taken not to record the result on a state-endorsed senior secondary certificate but for schools to issue a school completion statement of attainment in which employability skills achievement is recorded.

The construction of a portfolio can be described as a self-assessment activity; the items included in the portfolio are selected by candidates and reflect candidates' views about what is important and what constitutes appropriate evidence of their achievements. Of course, self-assessment includes approaches other than portfolio construction. Self-assessment is potentially valuable in promoting learning (Boud, 1995, 2002). It is represented in the model framework (see Figure 8) with Control of Assessment located with individual students.

		Degree of standardisation					
		Low			High		
Control over assessment	Centralised	National	Holistic judgment	Free assessment tasks	Common assessment tasks	Standardised tests	
					Prescribed tasks Prescribed standards Centrally marked	Externally set External standards Externally marked	
	Distributed	State		Recommended tasks Recommended standards Teacher marked	Prescribed tasks Prescribed standards Locally marked, but centrally or regionally moderated		
			School	Collective teacher judgment	Agreed tasks Agreed standards Teacher marked	Teacher set tasks Teacher selected standards Teacher marked	
				Teacher	Individual teacher judgment	Teacher-selected tasks Teacher-set standards Teacher marked	
Student	Self-assessment	Portfolio construction					

Figure 8: Framework for assessment models

Possible approaches to assessment

A literature review was conducted to identify approaches that have been applied to assessing generic employability skills in Australia and elsewhere. Very few examples were located that could be described as theoretically well informed, good practice according to the evaluation criteria that are applied in this project. Some examples from the review are included in Appendix 3. Other approaches were located from a review of the assessment of complex tasks in school subjects. The following approaches were identified as being potentially valuable for the assessment and reporting of employability skills achievement.

Table 6: Summary descriptions of possible assessment approaches

Approach	Description
Common assessment task	Tasks set centrally, for which a common set of performance levels are developed. Marking occurs within schools using the common criteria.
Embedded assessment	Existing subject assessment activities used, which provide evidence against the employability skills facets, teachers grade students' achievement of the employability skills.
Portfolio construction	Students construct portfolios of evidence of their demonstration of each employability skill.
Self-assessment	Using provided performance descriptors, students assess their performance in their choice of activities they have undertaken in both school and non-school settings.
Standardised assessment/testing	Standardised tests are developed using the facets of the ESF to define the construct domains for the test. These tests are administered under standard conditions (possibly online) to students. These tests are likely to include both multiple-choice and constructed response items.
Teacher-generated task	Teachers devise tasks that provide opportunities for the development of facets of employability skills and grade students' responses against suggested criteria.
Teacher-group assessment	Groups of teachers who have taught students meet and, using common criteria, make judgments about the level of performance by individual students on each of the employability skills.

Assessment approaches and employability skills

It is possible that no single assessment approach will adequately enable the assessment of all employability skills. Indeed, it may be the case that different facets of each employability skill would require different methods. It is desirable, however, in order to constrain the complexity of the assessment regime, to minimise the number of assessment approaches that are established. In order to test the number of potential assessment methods that may be needed, a clustering analysis was undertaken. One of the concerns that had arisen in discussion of employability skills is that some reflect substantially cognitive activities involving knowledge and skill, while others have a stronger affective component involving attitudes and dispositions. Further, some of the facets represent lower levels of knowledge such as recall and application while others demand selection and evaluation. Each employability skill's facets were classified along two dimensions, the first being a cognitive-affective dimension and the second a knowledge level one, from recall and application at the low end to metacognition at the high end of this factor. The mean score of the facets of each employability skill was plotted on these dimensions. In addition to facets of the eight employability skills, the personal attributes of the ESF were also analysed. This was done as a validity check, as they were expected to define a high end-point on the cognitive-affective dimension. There is no intention to include the personal attributes in the assessment proposals being developed in this project.

The results of this analysis are presented in Figure 9. The analysis reveals that two main clusters are apparent, and this suggests that no more than two assessment methods would be required to assess adequately all employability skills. Indeed, it may be possible that a single method would suffice. A third cluster, comprising basic and technology skills is also apparent.

The separation of basic skills from the other two groups of employability skills has two implications. First, some basic skills are being tested in what will become a national testing program from 2008. The final tests that students will undertake in the series will occur in Year 9. If these data are available, they could inform reports of students' achievement of the basic skills components of the ESF. If these data are available for reporting, assessment

efficiency suggests that these skills should not be re-assessed as part of an employability skills assessment regime. Further, if the basic skills component of the ESF is to be assessed as part of the national literacy and numeracy testing program, it would be desirable to include technology as a component of this program. Second, if the basic skills are assessed separately, the remaining facets of communication make this skill more affective and metacognitive than it is perceived now, and this may result in a preference for a different assessment approach than might be envisaged using the full range of facets.

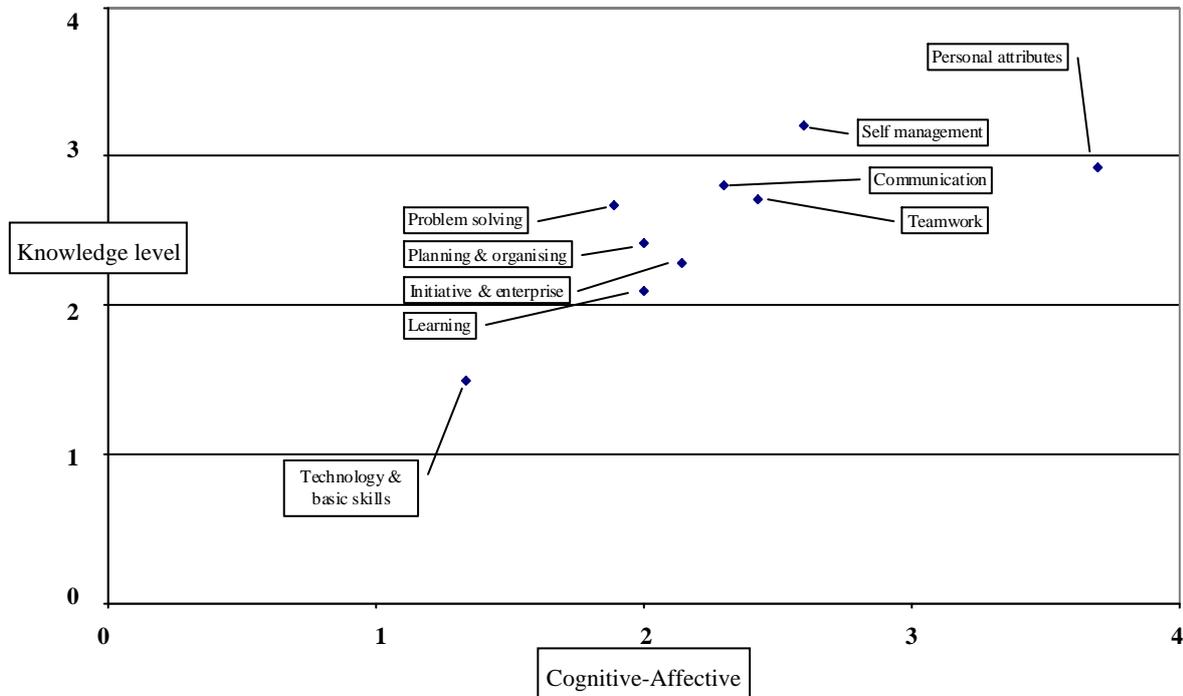


Figure 9: Clusters of employability skills along cognitive–affective and knowledge level dimensions

Application of criteria for evaluation assessment approaches

The criteria defined in Chapter 2 (Fundamental Concepts and Issues) were applied to each of the assessment approaches. The evaluations included in Table 7 were compiled from input from the research team and members of the Advisory Group.

The legend for the fitness-for-purpose column is:

PE = Program or system evaluation

MA = Measuring student achievement

AL = Assisting learning.

Table 7: Evaluation of assessment approaches

Assessment approach	Fitness for assessment purposes	Validity	Reliability	Objectivity	Feasibility	Usability	Authority	Backwash effects
Standardised testing	MA: High AL: Low PE: High	Pencil and paper tests are limited. Some elements of ES can be assessed well; others cannot. Low authenticity. Lacks practical component. Restricted validity.	High. Piloting of instruments and method of administration leads to reliable assessments of known precision.	High. All scripts are graded against specific criteria.	High costs in development. Moderate administrative and reporting costs. Moderate effort required within schools to supervise administration.	Simple summary performance bands against described standards. Easy to use and compare. Little detail to support improvement. Employers would prefer practical demonstration.	High. Results based on validated test instruments and processes could be warranted by state assessment boards.	Low feedback to students and teachers. Some system-level feedback on aggregate performance. Could inform curriculum change.
Common assessment task	MA: High AL: High PE: Moderate	High validity if tasks are selected to reflect target constructs and are piloted through scoring phase.	Moderate provided scoring rubrics are well developed and raters are trained to use them consistently.	Moderate–high. Common tasks with common rubrics and moderation could raise this to high. Permits comparison.	Moderate–high costs in developing common tasks. Low to moderate costs in assessing tasks. Moderation will increase costs. Practical. Requires targeted funding.	Summary statements against standards might provide more information than a grade or band. Useful for employers.	Moderate, possibly high, depending on scoring rubrics and moderation of assessment.	Moderate–high. Formative feedback can be provided to learners. Changes to classroom practices may be suggested.
School (teacher) selected assessment tasks	MA: High AL: High PE: Low–Moderate	High validity provided appropriate tasks are selected and scoring rubrics target ES. Would require PD for raters.	Modest. May be improved through the use of rubrics, but task variability will limit reliability.	Low between schools; moderate within schools.	Modest–low costs. Some PD desirable, but not a systemic high-cost activity.	Summary statements against standards might provide more information than a grade or band.	Moderate, possibly low. May be warranted at the school level. Values teacher judgment.	Moderate. Formative feedback can be provided to learners. Revisions to instructional practices may lead to improved performance.
Teacher-group judgment	MA: Moderate AL: High (assuming timely feedback) PE: Low	High, provided PD to promote a focus on core ES constructs.	Modest. Good within schools, but low between schools.	Low between schools, but high within schools.	Low costs. Some PD required, but not a systemic high-cost activity. Could be demanding of teacher time.	A level of performance summary (letter grade) against a description of that level would provide a summary of judged achievement.	Moderate, possibly low. May be warranted at the school level.	Moderate. If it is undertaken on several occasions, feedback could be provided to learners on ways of enhancing achievement.

Table 7 (continued): Evaluation of Assessment Approaches

Assessment approach	Fitness for assessment purposes	Validity	Reliability	Objectivity	Feasibility	Usability	Authority	Backwash effects
Embedded development/assessment	MA: Low AL: Moderate PE: Unsuitable	Modest. May be enhanced using ES-specific assessment rubrics. Authenticity depends on affordances of selected context.	Modest. Satisfactory within classes, but likely to be low between classes and schools.	Low. Could be enhanced with the provision of PD and curriculum support materials.	Very low costs, although provision of PD would increase the cost but enhance validity and reliability. Reporting could be a problem.	Letter grade summary of achievement, provided specific ES assessment is undertaken. Useful for comparison, but see Reliability comment.	Moderate–low. Likely to be assessed in-class and within other assessments.	Moderate–high.
Portfolio construction	MA: Unsuitable AL: Moderate PE: Unsuitable	Moderate, assuming guidelines are provided on how to construct the portfolio and what evidence to include.	Very low. If portfolios are assessed by teachers, reliability is low to moderate. Cloning of good portfolios is a threat to reliability.	Very low. Could be improved with teacher assessment, but this will impose a high load.	Low system costs in portfolio construction, but high load on teachers if assessment is required.	Detailed evidence, but unless the portfolio is assessed, no useful summary information.	Very low, unless the content is assessed by teachers.	Moderate for individuals; high if formally assessed; low for systems.
Self-assessment	MA: Unsuitable AL: Moderate PE: Unsuitable	Low, variable. Can be enhanced with quality support materials and self-assessment tools.	Low, but could be scaffolded by the use of self-assessment tools.	Very low, subjective.	Low system costs. Relatively easy to implement. Development of self-assessment tools would add to system costs.	No useful summary information.	Very low. Low likelihood of any weight being given to a self-assessed report by potential employers.	Moderate for individuals; very low for others.

Note: PE = Program or system evaluation; MA = Measuring student achievement; AL = Assisting learning

Reporting issues

A variety of comments were received about the purposes and authority of reports. In particular, two comments acknowledged reporting as a central purpose of any assessment:

The purpose of assessment is to construct a report that has credibility with employers.

[Reports] need to be easy to understand and have (sic) validity and reliability.

The key points being made are credibility, validity and clarity. Credibility will arise if the reported results are consistent with other related achievements and observations made by students, parents and employers and if the reported results are supported by evidence. That is, there needs to be a high level of transparency in the process of development, assessment and reporting.

Validity, often perceived as a characteristic of assessment, carries over into reporting and into interpretations made of reported achievement. The reports must be quite clear about what is being reported, and what is not. Some of the comments made by employers in consultations were that what is learned in school is not the same as what occurs in workplaces. For example, in the consultations undertaken by Costley et al. (2007) for this project, differences between the skills developed through learning in a school context and those required in work contexts were noted:

Communication based on history is not communication as required in industry. You can't use these curriculum subjects to teach employability skills.

Doing problem solving in chemistry is not the same as what industry wants.

What is taught in schools does not match the needs of our groups.

However, a rider to the third of these comments revealed a paradox:

Schools are there to educate students not teach them industry skills.

These comments suggest that some employers are sensitive to the problem that what is learned in a school context is not the same as that learned (and required) in a work context (see Resnick, 1987). However, much of school learning is applied in work contexts. The issue at stake here is the extent to which transfer of learning is required. One solution is to ensure that reports are based on what can be observed within school contexts and that is relevant to later work contexts and for schools to avoid assessing constructs that are clearly workplace-specific.

Irrespective of the approaches to assessment that are used to assess employability skills, schools will necessarily be restricted to observations in school (not necessarily classroom and curriculum) contexts. Thus, any inferences drawn about student achievement or performance from reported results must acknowledge the context in which the assessments were made.

This context effect poses a challenge. One of the criteria raised in discussion was the ‘fitness for purpose’ of assessment. This criterion extends to reporting. If the learning being reported is based on school contexts, its relevance for work contexts remains to be demonstrated. Such a demonstration will depend on showing that what is learned in school contexts can transfer to other domains.

Precisely what is to be reported must be decided. Reports based on the results of assessments can refer to performance and achievement, to attributes and dispositions, and to indicators of potential by referring to ability or capability. That is, a report may be ‘objective’ in that it presents the results of an assessment in a ‘raw’ form and says and seeks to imply nothing more than that the individual undertook an assessment in a particular context and received a particular score. Alternatively, a report may be interpretive and seek to draw inferences from particular assessment activities and results and infer likely future performances.

The detail of how the report is presented requires attention. In some assessments, results are reported on a scale so that a numeric record of achievement is offered. Alternatively, the report may present the results in achievement bands or in described categories of performance. In the latter case, a decision is required about the number of bands or categories that will be used in the reports. These decisions will depend upon the precision and granularity available from the particular assessment approaches that are implemented.

In consultations, concerns were raised about the authority with which reports are issued. Some approaches to assessment, for example national standardised testing, would enable a report to be issued by the agency responsible for the development and administration of the assessment, and the scoring and grading of achievement. Other forms of assessment, for example teacher-group judgments, may be valid and reliable within the school where the judgments were made and the report may have the imprimatur of the school. However, the warrant of a school-based report of employability skills depends on the knowledge of teachers conducting assessments.

Employers and human resource and recruitment specialists indicated satisfaction with school-authorized reports. They indicated a preference for brief reports including grades awarded by teachers following objective assessment, but with scope to refer to situations in which the employability skill was applied. Such brief but informative reports could be useful in interviews during which a potential employer could ‘drill down’ to ascertain the candidates’ understanding of the employability skills.

Conclusion

After an examination of approaches to assessing generic skills in Australia and elsewhere, seven categories of assessment approaches emerged as possibilities for assessing and reporting student achievement in the employability skills. These are:

1. Common assessment task
2. Embedded assessment
3. Portfolio construction
4. Self-assessment
5. Standardised assessment/testing
6. Teacher-generated task
7. Teacher-group assessment.

An examination of these and other assessment models suggests two key dimensions: standardisation of the assessment tasks and contexts; and locus of control of setting and marking assessments. Together, these lead to conclusions about the extent to which reported achievement is comparable and to varying opportunities for backwash effects (see Chapter 2).

The seven approaches were evaluated against the five criteria as elaborated in Chapter 2: validity, reliability, objectivity, feasibility and usability. While each of the methods had particular strengths and weaknesses, three methods—standardised testing, common assessment tasks and teacher-group judgment—were rated well on at least several of the evaluative criteria and provided a basis for valid, reliable and fair reporting of achievement. They were selected for more detailed examination (see Chapter 5). These methods, in addition to fulfilling satisfactorily the five evaluative criteria, have other potential benefits, including leading to higher levels of employability skills achievement of young people seeking to enter the labour market.

Chapter 5 – Preferred Methods for Assessing and Reporting Employability Skills

Seven approaches to assessing the employability skills were evaluated against five main criteria as recorded in Chapter 4. In addition to the application of these criteria to the seven approaches, other consequences of these assessment approaches were identified and discussed. In this chapter, the evaluation of assessment options is developed further, taking into account implications that arise from them.

Main strengths and weaknesses of assessment options

The major strengths and weaknesses of each assessment option are summarised, and further qualifications outlined in some cases.

Standardised testing

Clear advantages of this method are its reliability and objectivity of marking. A major disadvantage is its limitation to assessing only the more cognitive elements of the employability skills well. Also, timing issues could prove impractical. Developmental costs would be relatively high. It may be inexpensive to develop a test of a subset of cognitive facets, but it would be quite expensive to develop and administer tests of high-level cognitive and more strongly affective elements.

For a standardised test to be useful, it has to be a secure test, and therefore little feedback can be provided on student learning and item behaviour. Although sample items can be made available to inform teachers and students of the scope of the test, it remains unlikely that the testing would lead to desirable changes in other curriculum and assessment practices. It is unlikely, therefore, to lead to cohort-level improvements in employability skills achievement.

For the facets for which standardised testing is suitable, this method could produce nationally comparable grades and therefore support simple, informative summary reports. Further, because of the need for central administration of the test, results could be issued under the auspices of a central agency. This should give the report a high level of credibility, but its scope would be limited to those facets for which it is suitable. The reporting on other facets would require a complementary form of assessment.

Common assessment tasks

This option rates well on most criteria. Its effectiveness, however, depends upon the quality of the assessment tasks, which have to be constructed for administration under common conditions and marking according to commonly applied criteria. Common conditions do not necessarily refer to formal point-in-time examinations. The question of whether tasks should be subject-specific or transdisciplinary remains to be answered. If a subject-based approach were taken, many more tasks would be required thus increasing the developmental load. This

is offset by the likelihood of subject-specific tasks engendering greater acceptance by teachers and students at the senior secondary level and having more positive backwash effects on the development of employment-related skills in existing subjects. This, in turn, could lead to higher levels of employability skills achievement at the cohort level.

Specifications for scoring student performance are developed along with the construction of the common assessment tasks and, provided teachers possess (or rapidly acquire) expertise in judging the quality of student work against the level descriptors, reasonable consistency in grades can be expected. Comparability of results could be enhanced by having several teacher-assessors per task. Social moderation aimed at enhancing comparability would add to the cost of the method although it would be an effective built-in opportunity for development of teachers' assessment skills as well as of their understanding of the employability skills *per se*.

The performance-level descriptors developed for each assessment task provide a basis for simple summary reports of achievement. For un-moderated assessment (i.e. teachers decide on the grade without external validation), the warrant for reports lies with the school. This would also be the case if there were 'internal' moderation (i.e. teachers having the grades they assign validated by other teachers in the same school). For externally moderated assessment, the reports are issued under the auspices of curriculum/assessment agencies in the states and territories. The credibility of school-warranted reports, and hence the elimination of the need to establish moderation procedures, would depend on whether they met the requirements of employers.

School (or teacher)-selected (or -devised) assessment tasks

The all-encompassing title for this form of assessment (as in the section heading above) refers to the situation where teachers select student activities that provide evidence of achievement in the employability skill(s). In this sense the task is teacher-'devised'. Such tasks introduce additional variability into the assessment and reporting regime. Selecting tasks that are valid instances of the target constructs depends upon individual teachers' understanding of the constructs and of how they can be represented in the activities that they select as assessment opportunities. These activities might be located within existing subjects or could be designed around facets of the employability skills. A further source of variation lies in teachers' abilities to adapt generalised performance-level descriptors to composing scoring rubrics or standards schema for specific tasks. Similar variability can be expected when teachers apply those standards in judging the evidence within student responses. Thus, the validity and reliability of this method is likely to be lower than that of common assessment tasks. The feasibility of the method depends on the capacity of individual teachers to develop and assess the tasks. In particular, this method, in comparison with common assessment tasks, shifts the costs of development from a central authority to

individual teachers. The costs may be similar, but more difficult to quantify in the case of teacher selected tasks.

Teachers' willing engagement in devising or selecting tasks is likely to enhance students' achievement of the employability skills as teachers come to recognise opportunities for developing these skills through existing activities. This may result in cohort-level improvements in employability skills achievement.

The grades assigned to tasks provide a basis for summary reports. The warrant of these reports would be limited to schools who, presumably, endorse teachers' grades. It is unlikely that these judgments could have formal endorsement from external agencies.

Teacher-group judgment

The major strength of this option is its capacity to encompass almost all facets of the employability skills. Its validity is high provided teachers' attention is drawn to the most salient features of the target constructs, and this can be achieved through professional learning, some of which may occur through the judgment process itself as more experienced teachers mentor less experienced colleagues on judging panels. The limitations of the method relate to its lack of comparability between schools. A high level of internal reliability arises from the consensus approach to judgment implicit in the method, but this does not extend across schools. Moderation is not practicable because the evidence used by teachers is experiential and observational. Even if teacher panels within a school were asked to document the evidence on which their judgments were based, it is unlikely that this summary evidence would be an adequate substitute for the primary evidence, and therefore the moderated judgments would be less credible than the primary ones. Moreover, the load imposed by documenting evidence would add to the cost and therefore diminish the feasibility of the method.

Summary reports of achievement are available from this method. An individual student's grade arises from the agreement of a group of teachers of that student (teachers from different subject backgrounds) about that student's overall performance level on a particular employability skill (which would be manifest in different ways in different subjects). It is expected that the reports would carry the imprimatur of the school; it seems unlikely that such reports would be warranted beyond individual schools.

Embedded development and assessment

The major strength of this method lies in the potential for its backwash effects to be positive. It depends on curriculum mapping in which facets of employability skills are identified in existing syllabus statements and assessment guidelines. If the mapping exercise produces only a concordance of facets and subject content and if achievement of employability skills is inferred once the subject content has been mastered, there will be no backwash effect and the method would cease to be useful. If, on the other hand, the mapping is used to identify

and exploit opportunities for the development and assessment of employability skills within subject contexts, it could result in cohort-level improvements in employability skills achievement. The limitations of the method are that it depends upon individual teachers knowing the facets of the ESF and recognising their presence in their subject content and teaching and assessment processes. This aspect of the method introduces considerable variability as some teachers would be more engaged than others in recognising and exploiting curriculum opportunities.

This option is a variant of the method now endorsed in the VET sector. In the VET model, key elements of training packages—analogue of syllabus and assessment statements in the schools sector—incorporate the results of mappings (facets of employability skills being the assessable competencies). The mapping is undertaken centrally by developers of training packages. An important difference between the VET and schooling sectors is that qualifications in the VET sector have prescribed sets of units in which are embedded competency elements, including facets of employability skills. In senior secondary education in most jurisdictions, there is a single qualification that can be attained through the study of subjects selected from a vast range of subjects on offer (such that the number of subject combinations chosen over all students is very large). Under these circumstances it would be quite unrealistic to ensure that the curriculum experiences of all students exposed them to the development of all the employability skills. The application of this method would require that either every subject included all eight employability skills or that all possible combinations of subjects included them for, if not 100% of the cohort, for at least 95% of the cohort.

For a particular student in a particular year level at school, the method would produce a grade for each employability skill at the subject level. These grades would need to be collated by year-level coordinators. The warrant for reports would lie with the school, with little prospect of between-school moderation (comparability).

Portfolio construction

A distinction is drawn between portfolio construction and portfolio assessment. In constructing a portfolio, students select material that they believe provides evidence of achievement, including achievement of employability skills. This is a form of self-assessment, and one set of implications arises from this. If portfolios are subsequently assessed by teachers, or if teachers provide feedback on the portfolio, then the focus becomes the assessment of a product and the relevant issues are those that relate to all assessment options.

Portfolios have been recommended at some time in all three education sectors in Australia. Education.au trialled e-portfolios in the school sector and found some practical problems (Curyer, 2006). Because many students wish to include multimedia material in their

portfolios, the storage demands are large and the system becomes expensive. Further, if it is necessary to monitor what is stored, this adds to the system cost. Thus the feasibility of portfolios, specifically e-portfolios, is questionable. Their validity is questionable on some occasions: Is the student being rewarded for the quality of the work in the individual items in the portfolio or for organisational and presentation skills?—and, valuable as those skills are, they are not necessarily targeted constructs.

In the higher education sector, in addition to resource issues, questions have been raised about the imputed authority and responsibility of institutions for the nature of the content that students include in their portfolios. A specific concern relates to storage of content on an institutional website, with the implication that the institution endorses it. Endorsement of content can only occur if assessed by teachers, and assessing portfolios imposes a substantial load, thus reducing its feasibility as a method for assessing employability skills.

The portfolio's strength is that it provides, in some detail, supportive evidence of achievements that may be summarised in other reports. Thus, unless a portfolio is the primary evidence for assessment, the portfolio—paper-based or electronic—is an aggregation of supportive evidence rather than a report of performance.

Self-assessment

According to the prescribed evaluation criteria, self-assessment of employability skills has no particular strength. Its only benefit arises as a result of its backwash effects. Provided students are equipped to make useful assessments of their status on the employability skills, their own judgments may encourage them to seek opportunities to develop and document these skills. Enhancing students' capabilities for self-assessment is a necessary component of lifelong learning, a component of one of the employability skills (Boud, 1995; Sadler, 1989). It is also a necessary element of continuous improvement and, therefore, has value as an attribute of itself. As an assessment option in the present context, however, it would not achieve the intended purposes of providing a credible indicator of performance on employability skills or of providing the basis for acceptable reports.

Consequences of assessing and reporting employability skills achievement

The intended outcomes of assessing and reporting employability skills achievement are considered to be the reports of achievement, a greater awareness of the nature and importance of employability skills, and generally higher levels of these skills in the cohorts of young people completing senior schooling. Other consequences for consideration are now discussed.

Authority of reporting or certification

Reports of students' achievement of employability skills will have greater credibility if the development of assessment instruments and the assessment processes leading to results are transparent and fair. If the report has the status of a certificate, the results would have to be

monitored and verified by an authority external to the school. This source could be a board of studies, an assessment agency, or an individual school. It is less likely to be a national agency since constitutional responsibility for school education lies with the states and territories although the remit of any proposed national board might include nation-wide standards in reported results in some areas of learning.

If central agencies were to issue reports (or certificates) or to place their imprimatur on reports issued by schools, the authority over assessment and moderation processes would reside with them. Reports do not have the same official status as certificates.

Comparability

Comparability may be desired at national or state/territory levels. The only assessment method guaranteeing comparability on a national scale is standardised testing but, as noted elsewhere, this method is not sufficiently comprehensive and it gives primacy to reliability over validity. With moderation, common assessment tasks could provide comparability between jurisdictions but achieving national comparability would impose another layer of quality assurance (e.g. moderation, calibration etc.) Whether this level of comparability is warranted is a matter for judgment based on its perceived benefits and estimated costs.

Consultation

It is apparent that a round of consultations with curriculum/assessment agencies in all jurisdictions and school system representatives will be required, irrespective of what might be the preferred approach to assessing and reporting employability skills achievement.

The agreement of those stakeholders on the authority with which reports are issued will be needed as will agreement on associated arrangements; for example, mandatory assessment and reporting of employability skills and establishment of quality assurance processes such as moderation.

If standardised testing is selected, agreement will be needed as to the timing of the test(s) and the time available for testing.

Professional learning

One of the obstacles to the successful implementation of assessing and reporting on key competencies was a lack of investment in professional learning opportunities. This has been recognised as a need in relation to employability skills (Ratio Pty Ltd & Down, 2003).

Different approaches to this matter are suggested, depending upon the particular assessment approach selected. Standardised testing would require little professional learning, although if testing were undertaken without preparation or any follow-up, it is not likely to have a sustained influence on levels of skill within cohorts of school leavers. The implementation method suggested for common assessment tasks has an implicit professional learning approach embedded in it. A review of the recommended pilot program should evaluate the extent of teacher learning arising from the program. The teacher-group judgment method

would require a conventional professional development with three main thrusts— understanding the nature of the employability skills, understanding the nature of the evidence of achievement of the skills, and using protocols for reaching agreement within a group of teachers with respect to individual students.

The development of related skills

Employability skills are one group of a range of work-related knowledge and skills that have been proposed or implemented in schools on a national basis, including career education, enterprise skills and vocational skills, and vocational education and training in schools. In addition, other related skills development programs have been implemented within individual jurisdictions. These include essential learnings (in several states), capabilities (in South Australia), and repertoires of practice and ‘core skills’ (Queensland). To varying degrees, these skill complexes overlap MCEETYA’s (2003) endorsement of employability skills that recognised them as part of the total skills set required by young people.

Employability skills are developmental. At the point of exiting from school and entering work or further study, students do require these skills at some level, yet to be defined. But it must be acknowledged that these skills, in both the cognitive and affective domains, do begin to emerge from an early age. For example, communication skills, including a facility for negotiation, emerge long before children start school. And so a focus of the valuable skills for transition to work and further learning could be associated with other related learnings, especially career education, in the early years of secondary schooling. Two useful precursors to the assessment *of* employability skills in the senior secondary years would be students learning *about* them and gaining facility *with* them during their early secondary schooling. These antecedents could be realised by incorporating employability skills in career education. This approach would be consistent with MCEETYA’s (2005) recommendation to explore formative assessment approaches.

Viable approaches to assessing and reporting employability skills achievement

Of the seven assessment approaches evaluated above, three are regarded as acceptable according to the five main evaluation criteria of validity, reliability, objectivity, feasibility, and usability. The three approaches that remain in contention are:

1. Standardised testing;
2. Common assessment tasks; and
3. Teacher-group judgment.

Each of these approaches has strengths and weaknesses that are unique to it. Different approaches would be preferred if the evaluation criteria were weighted differentially.

Stakeholder perceptions

Our consultations suggest that employers would value the reliability offered by standardised testing but would prefer other—possibly supplementary—approaches in pursuit of validity,

to ensure development of the full range of employability skills and the application of them to practical situations.

School leaders expressed some concern about the impact on schools of having to schedule a common time for the standardised testing, but there seemed to be a view that if this were the only impact, it would be manageable. Some believed that standardised tests should not be a once-only activity because the employability skills are developmental. Repeated testing, however, would increase the assessment load and may require large scale item-banking, which would increase the developmental costs. Some questioned the capacity of teachers to make judgments about students' achievement of employability skills without further professional learning, but were reticent to allocate much time to that learning given the many other calls on teachers' time. Some felt that common assessment tasks were workable and foresaw some other advantages deriving from their use.

Representatives of parent organisations were concerned about the assessment load and believed that generic skills were being developed through existing subjects and co-curricular activities, but believed that employment-related skills embedded within subjects needed to be given greater prominence. Some believed that learning about employability skills could be linked to other aspects of career education, but believed that this linkage should be strengthened.

The major advantages of standardised testing are its reliability and perceived objectivity. It could be used to generate numerical scores of known precision. These scores could be reported on a nationally comparable basis and could appear as grades or performance bands on reports or certificates that are centrally produced.

The method would be most applicable to some foundational facets of communication, problem solving, planning and organising, and knowledge of learning. It is not likely to be applicable to important facets of teamwork or self-management. 'Speaking clearly and directly' is an example of a facet that would require a different form of assessment. If a single assessment approach is to be taken, standardised testing could not be that one approach. On the other hand, if a mixed-method approach were taken, standardised testing would suffice as one of the methods in the mixture (say, alongside or within common assessment tasks).

Common assessment tasks

With acceptable performance on all five evaluation criteria, this method is promising. It has the potential to be used for the foundation facets of all eight employability skills.

A matter requiring further consideration is the process by which an adequate number of assessment tasks could be developed and trialled. Given the vast number of senior secondary subjects across the various jurisdictions, it would be unrealistic to attempt to develop tasks for the eight employability skills for them all. However, by targeting curriculum areas, key

subjects in those areas—those with large enrolments—could be the focus of assessment task development and trialling. Task development would require setting situations, suggesting activities and objectives and constructing assessment guidelines so that grades can be awarded that are consistent with the performance-level descriptors presented in Chapter 3 (Performance Levels).

A second matter that requires further deliberation is the extent to which comparability is desired at regional, jurisdictional or national levels. The terms of reference for this study include ‘nationally comparable’ in a list of required properties of the proposed methods for assessment and reporting alongside valid, reliable, objective, feasible, and usable. It might be that it is sufficient for the method of assessment to be the same across the country and the decision making about student performance levels to be consistent across the country without seeking comparability of results in its purest form; that is, for student work of equivalent standard to be awarded the same grade. In order to achieve this sort of comparability, the establishment of increasingly complex moderation procedures would be needed. As breadth and complexity of procedures increases so do associated costs in both financial and human terms.

Teacher-group judgment

Teacher-group judgment has been shown to be a feasible method for reporting on student achievement of the generic and employability skills (McCurry & Bryce, 2000). For an assessment from a group of teachers to be a suitable basis for reporting achievement on all eight employability skills, teachers would have to be confident that the situations (student activities) observed in school settings (in different subjects) yielded the evidence to support judgments about the facet standards for each skill. The notion of a panel of expert judges is already in place (as in social moderation carried out within the school) so, on top of this, moderation across schools would not be feasible. Thus, reports on student achievement would be generated by schools without any external validation.

Possible implementation strategies

Implementation of the above methods will require different strategies.

Standardised testing

The foundational facets for each employability skill are the starting point for developing assessment/test instruments. These facets require some elaboration so that scoping of assessment constructs and development of range statements can proceed. From these statements, item specifications are written and the task of item development, using teams of item writers, undertaken. Items are reviewed by expert panels, trialled on a student population, and revised in the light of trial data. Item development is contracted to an organisation with demonstrated expertise in the area.

For employability testing to occur, the agreement of jurisdictions is required. It is unlikely that schools would agree to this testing occurring any later than first quarter of Year 12. In fact they might prefer for it to happen in Year 11 (and this timing complements the notion of 'entry-level' skills. Consultations about the timing of testing need to be undertaken with government, Catholic and independent school systems in each jurisdiction. A common time for test administration is required.

If employability skills achievement is to be recorded on existing certificates of education or if the report is to have the imprimatur of a curriculum/assessment agency, such agencies need to be consulted about the proposed reports.

Common assessment tasks

The development of common assessment tasks could be undertaken by many groups. There is considerable expertise in curriculum/assessment agencies and in teachers' professional (subject) associations.

Two related strategies are suggested. If the development of nationally consistent subjects is to proceed (i.e. English, Mathematics, Science and History learnings and/or standards are to be consistent across jurisdictions for the compulsory and non-compulsory years of schooling), it would be fitting to use these particular subjects as vehicles for carrying assessment tasks that emphasise foundational facets of employability skills. The tasks thus generated could even be required as part of the assessment regime for the subjects.

An alternative approach that does not depend on notions of national consistency is to establish a centrally coordinated project with support and review mechanisms attached to it, which vests in subject associations the responsibility for developing common tasks for assessing employability skills. Tenders could be sought for a pilot program for the development and trialling of a limited number of tasks. Subject associations, possibly in partnership with other organisations having general assessment expertise, could be targeted for these activities. If curriculum/assessment agencies are not directly involved in the pilot programs, they would need to be consulted about the design of the pilot program. The mapping undertaken by ACACA (2003) is an important source of data on employability skills already embedded in curriculum documents. Advantages of involving subject associations in the process of devising common assessment tasks include the likely embedding of these activities within the content and teaching processes of existing subjects, and the trialling of tasks developed through subject associations may have greater traction among teachers than externally developed tests. The trialling might therefore engender engagement of teachers and learners with the activities. The trialling might also contribute to the professional learning of teachers about foundational facets of employability skills and about assessment methods for complex constructs. An additional outcome of a pilot project

would be a collection of graded samples of student work. These work samples could become valuable resources for teachers who wish to compare their marking against exemplars.

It is also worth noting that having groups of teachers classified on the basis of their teaching subjects rather than on the state/territory where they live and work might break down some of the barriers to a national approach to some aspects of education in Australia.

On the assumption that pilot projects are successful, and using the learnings that arise from them, further assessment tasks could be developed for additional employability skills in other senior secondary subjects.

As for other methods, consultations need to be undertaken with school systems and curriculum/assessment agencies in each jurisdiction. The key issues on which agreement is required include: mandatory assessment and reporting, format of the reports, authority with which reports are released and, if applicable, the nature and extent of moderation required.

Teacher-group judgment

For this method to be implemented, agreement is required from the school systems in each jurisdiction. The school systems, at that level, need not have a role in undertaking or endorsing assessment results as the responsibility for conducting assessment and reporting in the employability skills resides with individual schools. In the case of mandated assessment and reporting, the school systems would indeed have a role.

Resources need to be developed and provided to schools. These resources include documents describing employability skills and their foundational facets, as well as procedures for making judgments about facet standards and the overall performance level (the so-called 2-stage process). Other resources include templates for recording judgments and summarising evidence provided, and professional learning opportunities for, say, one or two senior secondary teachers per school. The typical model (although others could be created) expects teachers to relay their understandings to colleagues within schools. The development of resources and professional learning opportunities would be based on successful responses to requests for tender.

Conclusion

Assessment of the employability skills that conforms to the hallmarks of educational measurement (e.g. validity and reliability) can be seen to be highly complex, placing a high—even unrealistic—level of demand on the education system. Rigorous assessment requirements may lead to the eventual abandonment of the goal as not worth the effort. The point should not be lost that a major achievement would be instilling employability skills into student learning so that transition from school to work is more efficient and effective from any perspective. Reporting achievement is important, and while a Rolls Royce may be the desirable vehicle, if it is unobtainable then some progress is at least possible with a Citroen 2CV—less than rigorous assessment may be the necessary price to pay for a general improvement in the employability skills of Australia's young people.

References

- ABS. (2007). *Schools Australia 2006* (Cat. No. 4221.0). Canberra: Australian Bureau of Statistics.
- Access Economics. (2005). Review of higher education outcome performance indicators. Canberra: DEST.
- Airasian, P. W. (1994). *The impact of the taxonomy on testing and evaluation*. In L. W. Anderson & L. A. Sosniak (Eds.), *Bloom's taxonomy. A forty year retrospective. Ninety-third yearbook of the National Society for the Study of Education. Part II* (pp. 82–102). Chicago: National Society for the Study of Education.
- Allen Consulting Group. (2004). *World-class skills for world-class industries: Accelerating reforms to vocational education and training*. Canberra: Australian Industry Group.
- Allen Consulting Group, & National Centre for Vocational Education Research. (2004). *Development of a strategy to support the universal recognition and recording of employability skills*. Directions Paper. Melbourne: Allen Consulting Group.
- Allen Consulting Group. (2006). *Assessment and reporting of employability skills in Training Packages*. Canberra: Department of Education Science and Training.
- American Educational Research Association, American Psychological Association, & National Council on Measurement in Education. (1999). *Standards for educational and psychological testing*. Washington, DC: American Educational Research Association.
- Anderson, J. R., Reder, L. M., & Simon, H. A. (1996). Situated learning and education. *Educational Researcher*, 25(4), 5–11.
- Annett, J. (1989). *Training in transferable skills*. Sheffield: Training Agency.
- Assessment Reform Group. (1999). *Assessment for learning: Beyond the black box*. Cambridge: University of Cambridge School of Education.
- Australasian Curriculum Assessment and Certification Authorities. (2003). *The employability skills framework report of mapping task undertaken by ACACA agencies: ACACA*.
- Australian Chamber of Commerce and Industry, & Business Council of Australia. (2002). *Employability Skills for the future*. Canberra: Department of Education, Science and Training.
- Australian Council for Educational Research. (2000). *Graduate Skills Assessment*. Test development and progress report. Camberwell, Vic.: ACER.
- Australian Council for Educational Research. (2001). *Graduate Skills Assessment*. Summary Report. Canberra: DETYA, Higher Education Division.
- Australian Education Council. (1989). *Common and Agreed National Goals for Schooling in Australia* (The Hobart Declaration on Schooling). Retrieved 10 October, 2000, from <http://www.mceetya.edu.au/mceetya/default.asp?id=11577>

- Australian Education Council. (1991). *Young people's participation in post-compulsory education and training*. Report of the Australian Education Council Review Committee. (Mr Brian Finn, Chair). Canberra: Australian Government Publishing Service.
- Australian Education Council. (1992). *Putting general education to work. The key competencies report*. (Mr Eric Mayer, Chair). Canberra: Australian Government Publishing Service.
- Australian Education Council. Mayer Committee. (1992). *Key competencies*. Report of the Committee to advise the Australian Education Council and Ministers of Vocational Education, Employment and Training on employment-related Key Competencies for post-compulsory education and training. Canberra: Australian Education Council and Ministers of Vocational Education, Employment, and Training.
- Australian Industry Group. (2004). *Skilling the existing workforce*. Canberra: Australian Industry Group.
- Australian National Training Authority. (2002). *Fresh thinking about learning and learners. A blue sky project*. Melbourne: Australian National Training Authority.
- Australian National Training Authority. (2003). *Employability skills in training packages*. Policy engagement forum, 28 May. Melbourne: Australian National Training Authority.
- Australian National Training Authority. (2004). *Training package development handbook*. Melbourne: Australian National Training Authority.
- Australian Parliament. House of Representatives Standing Committee on Education and Training. (2004). *Learning to Work*. Report on the inquiry into vocational education in schools. Canberra: Parliament of the Commonwealth of Australia.
- Biggs, J. B., & Moore, P. J. (1993). *The process of learning* (3rd ed.). New York: Prentice Hall.
- Boud, D. (1995). *Enhancing learning through self assessment*. London: Kogan Page.
- Boud, D. (2002). *Assessment: have we forgotten something?* In Australian National Training Authority (Ed.), *Fresh thinking about learning and learners: a blue sky project* (pp. 42-45). Melbourne: Australian National Training Authority.
- Bowden, J., Hart, G., King, B., Trigwell, K., & Watts, O. (2000). *Generic capabilities of ATN university graduates*. Retrieved 28 August, 2000, from <http://www.clt.uts.edu.au/ATN.grad.cap.project.index.html>
- Bryce, J, & McCurry, D. (2007). *Refinement of the employability skills for assessment at Year 12*. Discussion Paper prepared for DEST Advisory Group, Assessment and reporting of the employability skills of senior secondary students. ACER: Author.
- Burke, L, & Eraut, M. (1989). Cited in Jessup, G. (1990). *Common learning outcomes; Core skills in A/AS levels and NVQs*.
- Business Higher Education Round Table. (2003). *Developing generic skills: examples of best practice*. (B-HERT News 16, Special Issue). Melbourne: Business/Higher Education Round Table.

- Choppin, B. H. (1997). *Objective tests*. In J. P. Keeves (Ed.), *Educational research, methodology, and measurement: an international handbook* (pp. 771–775). Oxford: Pergamon.
- Claxton, G., & Carr, M. A. (2002). Tracking the development of learning dispositions. *Assessment in Education*, 9–37 (with peer commentaries by Lilian Katz, Frank Coffield, Royce Sadler).
- Clayton, B., Blom, K., Meyers, D., & Bateman, A. (2003). *Assessing and certifying generic skills. What's happening in VET?* (Research report No. NR1008). Adelaide: NCVER.
- Conference Board of Canada. (2000). *Employability skills toolkit for the self-managing learner*. Retrieved 04 February, 2001, from <http://www.conferenceboard.ca/education/learning-tools/toolkit.htm>
- Costley, D., Power, A., Watson, K., Steele, F., & Sproats, E. (2007). *Employers views of the employability skills of senior secondary school students*. Sydney: University of Western Sydney, College of Arts, School of Education.
- Crittenden, B. (1997). *The proper role of universities and some suggested reforms*. In J. Sharpham & G. Harman (Eds.), *Australia's future universities* (pp. 86–103). Armidale, NSW: University of New England Press.
- Curtis, D. D. (forthcoming). *VET pathways taken by school leavers* (LSAY Research Report). Melbourne: Australian Council for Educational Research.
- Curtis, D. D., & Denton, R. (2003). The authentic performance-based assessment of problem-solving. Adelaide: NCVER.
- Curtis, D. D., & McKenzie, P. (2002). *Employability skills for Australian industry: Literature review and framework development*. Canberra: Department of Education, Science and Training.
- Curyer, S. (2006). *The Australian experience: Developing an employability skills e-portfolio*. Paper presented at the 4th International E-portfolio Conference.
- Dawkins, J. (1988). *Higher education: A policy statement* (Ministerial Statement). Canberra: Australian Government Publishing Service.
- Deakin-Crick, R., Broadfoot, P., & Claxton, G. (2004). Developing an effective lifelong learning inventory. *Assessment in Education*, 11(3), 248–272.
- Earl, L. M. (2005). *Assessment as learning: Using classroom assessment to maximise student learning*. Thousand Oaks, CA: Corwin Press Inc.
- Erebus Consulting Partners. (2003). *The acquisition of employment related skills by secondary school students*. A final report to the MCEETYA Taskforce on Transition from School. Canberra: Department of Education, Science and Training.
- Field, L. (2001). *Employability skills required by Australian workplaces*. Sydney: Field Learning.
- Forster, M. (2005, October). *Reporting for learning, reporting as learning*. Paper presented at the 10th annual National Roundtable conference, Melbourne.
- Hager, P., Holland, S., & Beckett, D. (2002). *Enhancing the learning and employability of graduates: the role of generic skills*. Melbourne: The Business /Higher Education Round Table.

- Haines, C., Scott, K., & Lincoln, R. (2006). *Australian blueprint for career development*. Trial version. Perth, WA: Miles Morgan Pty Ltd.
- Higher Education Council, & National Board of Employment Education and Training. (1992). *Achieving quality*. Canberra: Australian Government Publishing Service.
- Hodgson, A., & Spours, K. (2000). *Qualifying for Success: Towards a framework of understanding*. London: Institute of Education.
- Lamb, S. P., & Vickers, M. (2006). *Variations in VET provision across Australian schools and their effects on student outcomes* (LSAY Research Report No. 48). Melbourne: Australian Council for Educational Research.
- Loacker, G. (Ed.). (2000). *Self assessment at Alverno College*. Milwaukee, Wisconsin: Alverno College Institute.
- Lokan, J. (Ed.). (1997). Describing learning. Implementation of curriculum profiles in Australian schools 1986-1996. Camberwell, Vic: ACER Press.
- Luke, A., Matters, G. N., Barrett, R., & Land, R. (2000). *New Basics Project Technical Paper*. Brisbane: Education Queensland.
- Masters, G. N., Forster, M., Matters, G. N., & Tognolini, J. (2006). *Australian Certificate of Education: Exploring a way forward*. Report of the Australian Council for Educational Research for the Department of Education, Science and Training. Canberra: Australian Government.
- Matters, G. N. (1999). *Setting, describing, and presenting standards*. Unpublished manuscript. Brisbane: Queensland Board of Studies.
- Matters, G. N. (2005, September). *Designing assessment tasks for deep thinking*. Paper presented at the Curriculum Corporation Conference, Brisbane.
- Mayer, E. (1992). (Chair). Employment-related key competencies: A proposal for consultation. Melbourne: Mayer Committee.
- McCurry, D., & Bryce, J. (1997). *The school-based key competencies levels assessment project*. Final report. Canberra: DEETYA.
- McCurry, D., & Bryce, J. (2000). Victorian Board of Studies. *Key Competencies levels assessment trial. Working paper 2*. Melbourne: Victorian Curriculum and Assessment Authority.
- McCurry, D. & Bryce, J. (2007). *Identifying performance levels*. Discussion Paper prepared for DEST Advisory Group, Assessment and reporting of the employability skills of senior secondary students. ACER: Author.
- MCEETYA Transition from Schools Taskforce. Working Group on Key Competencies. (1996). *Key competencies for work, education and life*. Report on the outcomes of the key competencies pilot phase (No. 12/96). Canberra: DEETYA.
- MCEETYA. (2003). *Response to advice on employability skills*, MCEETYA Out-of-session resolution, December.

- MCEETYA. (2005). Note in response to the Report of the Transition from School Taskforce, 18th MCEETYA Meeting.
- McGaw, B. (2006). *Assessment fit for purpose*. Paper presented at the 32nd IAEA Annual Conference. Assessment in an Era of Rapid Change: Innovations and Best Practices. from <http://www.iaea2006.seab.gov.sg/conference/download/papers/Keynote-Assessment%20fit%20for%20purpose.pdf>.
- McLeish, A. (2002). *Employability skills for Australian small and medium sized enterprises*. Canberra: Department of Education Science and Training.
- McMahon, M., Patton, W., & Tatham, P. (2003). *Managing life, learning and work in the 21st century*. Issues informing the design of an Australian Blueprint for Career Development. Perth, WA: Miles Morgan.
- Messick, S. J. (1995). *Validity of psychological assessment*. Validation of inferences from persons' responses and performances as scientific inquiry into score meaning. *American Psychologist*, 50(9), 741–749.
- Messick, S. J. (1998). *Alternative modes of assessment, uniform standards of validity*. In M. D. Hake (Ed.), *Beyond multiple-choice: Evaluating alternatives to traditional testing for selection* (pp. 59–74). Mahwah, NJ: Lawrence Erlbaum.
- Milne, C. (2000). *Tertiary literacies: integrating generic skills into the curriculum*. In S. Fallows & C. Steven (Eds.), *Integrating key skills in higher education. Employability, transferable skills and learning for life* (pp. 87–97). London: Kogan Page.
- Ministerial Council on Education Employment Training and Youth Affairs. (1999, July). *The Adelaide Declaration on National Goals for Schooling in the Twenty-first Century*. Retrieved 10 October, 2000, from <http://www.mceetya.edu.au/mceetya/nationalgoals/natgoals.htm>
- Ministerial Council on Education Employment Training and Youth Affairs. (2004). National data on participation in VET in schools programs and school-based new apprenticeships for the 2004 school year. Canberra: MCEETYA.
- Moss, P. A. (1994). *Can there be validity without reliability?* *Educational Researcher*, 23(2), 5–12.
- Murray, T. S. (2003). Reflections on international competence assessments. In D. S. Rychen & L. H. Salganik (Eds.), *Key competencies for a successful life and a well-functioning society* (pp. 135–159). Göttingen: Hogrefe & Huber.
- National Curriculum Council. (1990). *Discussion Document on Core Skills*. York: Author.
- National Industry Education Forum. (2000). *The Key Competencies portfolio approach – a kit*. Canberra: Department of Education, Training and Youth Affairs.
- National Training Quality Council. (2005). Resolution: Final advice on the incorporation of employability skills in training package qualifications and related issues. Melbourne: National Training Quality Council.

- Oliver, K. M., Russell, C., Gilli, L. M., Hughes, R. A., Schuder, T., Brown, J. L., et al. (1997). *Skills for Success in Maryland: Beyond workplace readiness*. In H. F. O'Neil (Ed.), *Workforce readiness. Competencies and assessment* (pp. 47–76). Mahwah, NJ: Lawrence Erlbaum.
- Pellegrino, J., Chudowsky, N., & Glaser, R. (Eds.). (2001). *Knowing what students know. The science and design of educational assessment*. A report of the National Research Council. Washington, DC: National Academy Press.
- Pitman, J. A., Matters, G. N., & Nuyen, N. A. (1998, May). *A case for testing generic skills*. Paper presented at 24th annual conference of the International Association for Educational Assessment, Barbados, West Indies.
- Pitman, J. A., O'Brien, J. E., & McCollow, J. E. (1999, May). *High-quality assessment: We are what we believe and do*. Paper presented at 25th annual conference of the International Association for Educational Assessment, Bled, Slovenia.
- Powell, R., Smith, R., & Reakes, A. (2003). *Basic skills and key skills: A review of international literature*. Final report. London: NFER.
- Precision Consultancy. (2007). *Graduate employability skills*. Canberra: DEST, BIHECC.
- Quality of Education Review Committee. (1985). *Quality of education in Australia*. Report of the Review Committee. (Prof Peter Karmel, Chair). Canberra: Australian Government Publishing Service.
- Queensland Department of Education. (2003). *New Basics Project*. New Basics research program. Brisbane.
- Ratio Pty Ltd, & Down, C. (2003). *Employability skills in training packages*. Final report. Melbourne: Australian National Training Authority.
- Resnick, L. B. (1987). *Learning in school and out*. *Educational Researcher*, 16(9), 13–20.
- Ryan, C. (1997). *NSW key competencies pilot project report*. Sydney: NSW Department of Education and Training Coordination.
- Sadler, D. R. (1986). *Subjectivity, objectivity, and teachers' qualitative judgments* (Discussion Paper No. 5). Brisbane: Queensland Board of Secondary School Studies.
- Sadler, D. R. (1989). *Formative assessment and the design of instructional systems*. *Instructional Science*, 18(2), 119–144.
- Sadler, D. R. (2002). *Learning dispositions: Can we really assess them?* *Assessment in Education: Principles, Policy & Practice*, 9(1), 45–51.
- Shavelson, R. J., Gao, X., & Baxter, G. P. (1993). *Sampling variability of performance assessments* (CSE Technical Report No. 361). Los Angeles: University of California, CRESST.
- Stobart, G. (2004). *The formative use of summative assessment: Possibilities and limits*. Paper presented at 30th annual conference of the International Association for Educational Assessment, Philadelphia.

- Thomson, S., & de Bortoli, L. (2007). *Exploring scientific literacy: How Australia measures up*. The PISA 2006 survey of students' scientific, reading and mathematical literacy skills. Melbourne: ACER Press.
- Troper, J., & Smith, C. (1997). *Workplace readiness portfolios*. In H. F. O'Neil (Ed.), *Workforce readiness. Competencies and assessment* (pp. 357–382). Mahwah, NJ: Lawrence Erlbaum.
- Turner, D. (2002). *Employability skills development in the United Kingdom*. Adelaide: National Centre for Vocational Education Research.
- University of South Australia. (2000, December, 2000). Graduate qualities. Retrieved 21 March, 2001, from <http://www.unisanet.unisa.edu.au/extracourseinformation/gradquals.asp>
- Wiggins, G. P. (1998). *Educative assessment. Designing assessment to inform and improve student performance*. San Francisco: Jossey-Bass.

Appendix 1: Generic, work-related skills – a collation from different countries

AUSTRALIA		USA	UK	SOUTH AFRICA	CANADA
<p>Key Competencies</p> <p>Collecting, analysing and organising ideas and information</p> <p>Communicating ideas and information</p> <p>Planning and organising activities</p> <p>Working with others and in teams</p> <p>Using mathematical ideas and techniques</p> <p>Solving problems</p> <p>Using technology</p> <p>(Cultural understanding)</p> <p>Core Skills (Queensland)</p> <p>Comprehend and collect</p> <p>Structure and sequence</p> <p>Analyse, assess and conclude</p> <p>Create and present</p> <p>Apply techniques and procedures</p> <p>Employability Skills Framework</p> <p>Communication</p> <p>Technology</p> <p>Problem-solving</p> <p>Teamwork</p> <p>Initiative and enterprise</p> <p>Planning and organisation</p> <p>Self-management</p> <p>Learning</p>	<p>Clusters of Practice (New Basics)</p> <p>Live in and prepare for diverse family relationships</p> <p>Collaborate with peers and others</p> <p>Maintain health and care of self</p> <p>Learn about and prepare for new worlds of work</p> <p>Develop initiative and enterprise</p> <p>Blend traditional and new communications media</p> <p>Make creative judgments and engage in performance</p> <p>Communicate using languages and intercultural understandings</p> <p>Master literacy and numeracy</p> <p>Interact within local and global communities</p> <p>Operate within shifting cultural identities</p> <p>Understand local and global economic forces</p> <p>Understand the historic foundation of social movements and civic institutions</p> <p>Develop a scientific understanding of the world</p> <p>Work with design and engineering technologies</p> <p>Build and sustain environments</p>	<p>Scans Workplace Know-How</p> <p><i>Foundation Skills</i></p> <p>Basic skills (reading etc.)</p> <p>Thinking skills</p> <p>Personal qualities</p> <p><i>Competencies</i></p> <p>Resources (time allocation etc.)</p> <p>Interpersonal</p> <p>Information (acquires & evaluates information etc.)</p> <p>Systems (understands systems etc.)</p> <p>Technology</p> <p>National Committee of Inquiry into Higher Education</p> <p>The knowledge and understanding that a student will be expected to have upon completion of a program</p> <p><i>Key skills</i></p> <p>Communication, numeracy, the use of information technology, and learning how to learn</p> <p><i>Cognitive skills</i></p> <p>Understanding of methodologies or ability in critical analysis and so on</p>	<p>Core Skills</p> <p>Communication</p> <p>Personal skills</p> <p>Numeracy</p> <p>Information technology</p> <p>Problem solving</p> <p>Competence in modern language</p> <p>QCA, England</p> <p>Taking part in discussions and making presentations</p> <p>Reading and responding to written material</p> <p>Producing written material</p> <p><i>Information Technology</i></p> <p>Preparing information</p> <p>Processing and presenting information</p> <p>Reviewing the use of information technology</p> <p><i>Application of Number</i></p> <p>Collecting and recording data</p> <p>Working with data</p> <p>Presenting findings</p> <p><i>Working with Others</i></p> <p>Planning activities</p> <p>Working towards identified targets</p> <p><i>Improving Own Learning & Performance</i></p> <p>Setting targets and action planning</p> <p>Following plan to meet target</p> <p>Problem solving</p>	<p>Critical Cross-Field Outcomes</p> <p>Identify and solve problems in ways which display that responsible decisions using critical and creative thinking have been made</p> <p>Work effectively with others as a member of a team, group, organisation, community</p> <p>Organise and manage oneself and one's activities responsibly and effectively</p> <p>Collect, analyse, organise and critically evaluate information</p> <p>Communicate effectively using visual, mathematical and/or language skills in the modes of oral or written presentation</p> <p>Use science and technology effectively and critically, showing responsibility towards the environment and others</p> <p>Demonstrate an understanding of the world as a set of related systems by recognising that problem solving contexts do not exist in isolation</p>	<p>Employability Skills Profile</p> <p>Fundamental Skills</p> <p>Communicate</p> <p>Manage information</p> <p>Use number</p> <p>Think & solve problems</p> <p>Personal & Management Skills</p> <p>Demonstrate positive attitudes & behaviours</p> <p>Be responsible</p> <p>Be adaptable</p> <p>Learn continuously</p> <p>Work safely</p> <p>Teamwork Skills</p> <p>Work with others</p> <p>Participate in projects & tasks</p> <hr/> <p>OECD</p> <p>DeSeCo Key Competencies</p> <p>Acting autonomously and reflectively</p> <p>Using tools interactively</p> <p>Joining and functioning in heterogeneous groups</p>

Appendix 2: Analysis of merits of assessment options – stakeholder perspectives

		Standardised testing	Common Assessment Tasks	Teacher group judgment
Students and parents	Benefits	The testing routine is simple, familiar and minimally invasive. If the assessment is completed online, and provided items do not require manual marking, reports could be generated quickly.	Tasks can be designed and selected to be integrated within many existing courses. This may increase the perceived salience of these skills for students. A range of tasks that reflect each employability skill will lead to repeated exposure to the skills and therefore to their development.	Because the judgments of a group of teachers are based on extended observations of students in a range of school settings, students do not need to prepare for one or several specific assessment events. Thus the assessment is non-invasive. Nonetheless, students and parents would need to be informed at an early stage about the assessment method and the evidence being considered.
	Weaknesses	Because not all employability skills can be assessed through standardised tests, supplementary sources of evidence will be required for them. This will add to the complexity of the assessment regime and the load on students. Because standardised testing is likely to occur once, students will not have opportunities to develop their employability skills.	CATs will require students to invest time in the assessment. All learning requires the investment of mental effort and time. At issue is the extent to which enhanced skill levels justify the commitment.	If the judgment occurs once and near the end of schooling, students will not have an opportunity to develop areas where the need for improvement is indicated. This could be overcome by having two or more judgment activities, with specific feedback following the first of them. Students and parents would need to be assured that the focus of the judgment is on employability skills specifically and not other aspects of student-teacher interactions.
Teachers	Benefits	The testing routine is simple, familiar and minimally invasive.	Teachers will be directly involved in promoting employability skills and in their assessment. Through these activities, they should be alert to opportunities for their development within existing subjects.	While the method requires some organisation in bringing teachers together for this purpose, it does not impose a high burden on teachers to learn new assessment activities and methods.
	Weaknesses	The routine assessment may limit teachers' ability to provide employability skills instruction. Indeed, some teachers may not see a need to do this, as the assessment is managed externally.	Teachers will need to become familiar with a suite of assessment tasks and in particular with the assessment process and performance standards for facets. This will require effort and time.	Teachers will need to become familiar with the assessed facets and with the standards against which criterion judgments are required. While this will take some time, it is not regarded as an unreasonable burden.
Employers and further and higher education providers	Benefits	Achievement assessed using standardised tests are likely to be fine-grained, enabling ready discrimination between candidates. Reported achievement, based on standardised tests should have high credibility.	Employers should be confident that students will be very familiar with the employability skills and be able to describe examples of their application.	Employers should be confident that, within schools, teachers' judgments will be consistent and sufficiently fine-grained to enable comparisons.
	Weaknesses	The need to integrate results from other assessment methods may limit the precision of overall achievement results.	In the absence of moderation procedures, some variation in the application of standards can be anticipated. This variation is likely to be acceptable.	Reports based on judgments of students from different schools are not likely to be readily comparable.

Appendix 3: Generic and employability skills – assessment examples

Examples of approaches to the assessment and reporting of employability (and similar) skills are presented in this appendix. The examples are far from exhaustive. The assessment approaches in some examples do not lead to formal reporting. The examples are drawn from Australian and international experience and from schools, vocational education and higher education. The relevance of these examples to assessing and reporting the employability skills achievement of senior secondary students in Australian schools is discussed.

Standardised tests

Standardised testing of complex constructs requires the development of an assessment instrument, the administration of that instrument and scoring and grading candidate responses. Instrument development is a multi-stage process in which: the target constructs are defined conceptually; the definitions are operationalised; specifications for the test and test items, including response formats, are developed; draft items are written then subject to expert review; items are piloted; the instrument is revised and a final version of the test, including documentation, is produced (Gable & Wolf, 1993; Rust & Golombok, 1999). Tests are administered under prescribed conditions; scripts are marked; data are analysed; and student results are reported.

Extensive expert judgment is exercised in the development of the instrument and in producing the rules under which responses will be scored. Marking of scripts is often an automated process.

Graduate Skills Assessment

The GSA was developed through extensive consultations with university staff and employers about what characteristics of graduates those staff believed to be the most important. A list of 17 skills and attributes was identified through these consultations (ACER 2001, p. 27). In order to limit the scope of the trial of the instrument, four of these attributes were selected for development, namely communication (argument and report and writing), problem solving, interpersonal understanding and critical thinking. These constructs were ranked as the most important by university staff and the first three were also the top three among employer rankings. The employer's fourth ranked skill was analytic thinking, and it was regarded as being sufficiently close to critical thinking to justify the inclusion of that construct. It is worth noting that the 'big three' generic skills¹² of communication, teamwork (interpersonal understanding) and problem solving were identified as the highest priorities among university staff and employers.

¹² These three skills appear almost universally in generic and employability skills schemes in Australia and overseas.

It was intended to develop the scope of the instrument further by adding some of the other identified attributes including basic skills, management skills, information technology skills, research skills and additional items to discriminate among high achievers. There was also an intention to develop the test for online delivery. This would improve the cost-effectiveness of the testing. Thus far, these developments have not occurred. The test is secure, but sample items are made available to candidates so they are familiar with the format of the test and the types of questions that will be asked (ACER 2003). In this way, no candidate can be advantaged by special coaching as all students have access to the same information.

In several phases, the test instrument was administered to several thousand students from 19 Australian universities and to a sample of 400 senior secondary students. The three scales based on multiple-choice formats (problem solving, critical thinking and interpersonal understandings) had internal scale reliabilities in excess of 0.8, indicating that the scales were coherent representations of the constructs. Some differences in scores on the various scales were identified by sex, with females performing better than males on the written communication component but less well on the problem solving items. Differences were also detected by field of study. These differences are likely to reflect differences in general cognitive abilities, indicated by the range of tertiary entrance scores in various fields of study. Some of the differences in the relative strengths are likely also to reflect personal attributes associated with career preferences. For example, the relatively good performance of nursing students on the interpersonal understandings scale is consistent with the choice of this career by people with well developed interpersonal skills.

In summary, the GSA has good psychometric properties. The model has the potential to be developed to include other constructs and the technology exists to make the test available online. Scoring the written elements of the test (the report and argument writing tasks) is one of the more expensive aspects of the testing. The GSA model could be developed to assess the employability skills of senior secondary students. Additional scales will need to be developed to assess all eight employability skills. (See the summary of generic employability skills in Table 1.)

PISA Problem Solving Assessment

PISA is a major cross-national program in which students near the end of their compulsory schooling (15-year-olds) are tested on a range of 'literacies'. The tests are repeated in a three-year cycle, with tests having been run in 2000, 2003 and 2006. Students are tested on reading comprehension, mathematical literacy, scientific literacy and problem solving. A key point is that the testing is less focused on specific school curriculum content and more on students' abilities to understand, apply, interpret and draw inferences from given information. There are alternative testing programs that do concentrate on the achievement of curriculum content. There is, no doubt, a strong correlation between mathematical literacy

as measured in PISA and achievement on curriculum-based tests, since a learner who is capable of reasoning with numeric data will use that ability in their school mathematics tests. The PISA tests of problem solving are the particular focus of this review because the problems are designed to test students' abilities to apprehend and understand problems, to identify relevant given information, to apply that information, to reason deductively and inductively with it, to reflect on tentative solutions and to communicate their findings. The development of problem solving as a construct and of problem tasks that exercise the range of problem solving processes are described in *Problem solving for tomorrow's world* (OECD, 2004, pp. 25–39). There, sample problems are presented and their relationships to the component processes and the cognitive demands of each are explained.

In advance of the testing, three levels of student performance were anticipated. These levels are described as **standards** that characterise students' approaches to problem solving, given their observed behaviour on the assessment tasks. Level 3 problem solvers were able to identify relevant information in the task description, to reason about it, to identify possible solutions, to reflect on those solutions and make judgments about them, and to communicate their results. These learners are described as 'reflective, communicative problem solvers.' Learners at Level 2 are able to identify information and reason using it, but are less likely to monitor their solution attempts, reflect on them and communicate their solutions clearly. They are described as 'reasoning, decision making problem solvers.' At Level 1, 'basic' problem solvers are able to identify given information and to use it, but are unlikely to draw inferences from it. In the cross-national study, a proportion of young people were identified as operating below Level 1.

For each problem task that was developed, an analysis of the task revealed the processes that would be involved in its successful completion. Further, the shortcomings in unsuccessful solution attempts were diagnosed. Thus, the skills that differentiated successful and unsuccessful problem solvers were identified. Scoring keys were developed for each problem, so that various incomplete or inaccurate solutions could be given partial credit for the skills that were evinced.

This approach to task design and scoring, based on a principled analysis of the skills that are implicit in problem solving, enables well-targeted assessment and informative feedback. Although the primary purpose of PISA is national comparisons, individual learners can be informed about their level of problem solving. Further, the assessment scoring process is transparent, so that learners, given this feedback, could be guided towards the next level in their quest for improvement.

Because the primary purposes of the PISA testing are international (and jurisdictional) comparisons and comparisons among population sub-groups, the results of large numbers of students are aggregated to provide estimates of national and sub-group means. It is not necessary that the scores assigned to individual students are very precise, since the results of

students within defined groups are averaged. In reporting individual achievement, especially for high stakes purposes such as selection and sorting by potential employers, it will be necessary to ensure that the precision of individual measures is consistent with these purposes. Thus, if items such as those developed for the PISA testing are used for employability skills assessments, each candidate will need to respond to a greater number of items than they do in the PISA assessment.

The standards-based approach to defining performance levels is applicable to the assessment of employability skills in the schools sector. A particular advantage of the set of performance standards defined for problem solving in PISA is that a teacher could provide feedback to a student who had been graded at Level 1 or Level 2 and advise her about what she would need to do to improve her performance and achieve the next level. Such a practice would lead to improved aggregate achievement of employability skills within cohorts of young people completing senior secondary schooling.

The PISA model of problem solving assessment is applicable to some other employability skills.

Employability Skills Profiler¹³

The Employability Skills Profiler (ESP), funded by DEWR, was developed and trialled by Chandler MacLeod during 2006 and is being implemented during 2007. Its purpose is to assess the employability skills of unemployed persons who are clients of Jobs Network and Disability Employment Network service providers, to develop a profile of clients' employability skills, and to match their profile to the skills profiles of jobs.

The Employability Skills Profiler (ESP) objectively assesses a job seeker's generic or transferable skills and shows how well the job seeker's skills fit with the skills required by over 1000 job types. (DEWR, 2007)

The tool was described in some detail by Curtis and Grant as 'a tool for the objective assessment of employability skills'. It has been used to measure both the skill levels of individuals and the skill requirements of jobs, and therefore, to match individuals to jobs.

Chandler MacLeod analysed the employability skills and concluded that the eight skills could not be assessed coherently but that they could be represented as 36 constructs. The 36 constructs were mapped onto a variety of ability and personality measures for which existing

¹³ Other than those sources specifically acknowledged, the information presented on the ESP is derived from two sources. One was a paper presented by Rob Curtis and Lieschen Grant (both DEWR staff members) at the 2005 NCVER Research Conference at Wodonga. The paper was not published, but extensive notes were taken by David Curtis at the conference. The other source was a telephone conversation on 2 August 2007 with Mr Kevin Chandler, Executive Director of Chandler MacLeod.

instruments were available. It should be noted that ability and personality instruments treat these constructs as individual traits that are relatively stable over time. This is inconsistent with a view of employability skills as elements of human capital that can be and need to be augmented.

Subject matter experts identified the 36 employability skill elements required for more than 1100 jobs. The skills identified were classified as either required or desirable and a system of weighting was used to assess the level of performance level required for each skill. Five performance levels were identified for each component.

The profile of skills established through the administration of the ESP is being used to match people to the requirements of positions. It is used to enable individuals to identify skills strengths and skills gaps and to enable employers to match people to jobs.

Because of the match between individuals and job requirements, it enables the fit with the preferred jobs to be assessed, and therefore can be used as a career guidance tool taking into account local skills shortages. The ESP can be customised by employers to meet the specific skills needs of particular job roles. It can be used to develop an individual training program to improve job and person fit.

In addition to the job-match report, the ESP also results in a paper report for individuals focusing on individuals' strengths and was intended to be 'a feel-good report'.

The proposed matching of person and jobs highlights one of the definitional issues raised in the brief discussion above about what is meant by a skill and by the degree of abstraction with which generic employability skills are described. The skills requirements of particular jobs are, by definition, job-specific. In the set of skills identified for a particular position, it is possible to distinguish those skills that are highly job-specific and those that are common to many jobs. Those that are common can be classified into groups corresponding to the eight employability skills. However, when the job-specific skills are removed and a profile of the employability skills is developed, any given profile would apply to many jobs. Thus a claim for the utility of the ESP as a tool both for matching unemployed persons to jobs and for reporting on the achievement of generic skills must be questioned.

The online methodology and the ability to generate reports quickly from the system are advantageous. The output of such testing could be valuable as a diagnostic tool. A disadvantage of the system is that it uses standardised ability and personality instruments. The eight employability skills were mapped onto elements of existing psychological tests that assess what are presumed to be fixed traits. This represents a re-definition of the employability skills. The assumption that they are fixed traits does not augur well for their development.

Key Skills Certificate

A set of key skills was defined in the United Kingdom. The skills were:

- Communication
- Problem solving
- Working with others
- Application of number
- Information technology
- Improving one's learning and performance.

A Key Skills Qualification for 16–19-year-olds was introduced into England and Wales in 2000. This qualification targeted communication, application of number and information technology skills—the basic key skills. Those skills not covered by this qualification were re-labelled 'wider key skills'. Initially, all 16–19-year-olds in any form of education and training were expected to gain this qualification. The qualification was voluntary, although there were perceptions that it was compulsory and this led to an initial uptake that subsequently waned (Hodgson & Spours, 2000). By 2002, few of the students enrolled in academic qualifications, mainly at selective schools, participated—presumably because young people enrolled in academic tracks were regarded as competent in these skills and these basic skills were being developed through the academic programs. Uptake of the program was greater in colleges offering vocational qualifications (Powell, Smith, & Reakes, 2003), and this differentiation appears to have led to the qualification being seen as remedial rather than as an affirmation of employment-related skills.

It should be noted that the policy context of the UK was rather different from Australia's. Two reasons for introducing the Key Skills qualification were to broaden the upper level secondary curriculum, in which students often took only three A Level subjects, and to bridge an apparent divide between academic and vocational study (Hodgson & Spours, 2000). If the qualification has had an effect, it may have been to exacerbate that divide.

Students who completed the qualification at Level 3 were awarded 'a very generous' (Powell et al., 2003) 60 points towards their university entrance score (UCAS). This encouraged students in academic programs to undertake the Key Skills qualification, but not all universities use the tertiary entrance score—the more selective universities being less likely than others to recognise it. Thus, students and their parents have received mixed messages about the value of the qualification.

Hodgson and Spours (2000: 20) summarised the problems of the Key Skills Certificate as having 'too many complication, too little currency and [being] too difficult to achieve'. Turner (2002: 15–16) identified two main concerns with the key skills qualification. He noted 'the unsatisfactory experience of both educators and employers regarding the

assessment of the three basic key skills as they were developed into a national qualification. Testing procedures became too complicated and there is real concern that [the qualification had become] assessment and not learning dominated'. The central problem appears to have been the complexity of the moderation processes of tests of the basic key skills and portfolio evidence of these skills leading to the award of the qualification. Employers regarded the basic key skills as of lower importance than the wider key skills that were not assessed and were hesitant 'to place great store on a "certificate" or even a portfolio of evidence as proof of "having these skills"' (Turner, 2002: 17). To address the lack of breadth, Hodgson and Spours (2000: 30) suggested that the 'wider key skills' should become the focus of the qualification.

In summary, the Key Skills qualification was perceived to be too narrow and a test of basic skills only. This added little to what employers and higher education providers would know about candidates based on other school achievement information. The qualification was complex, in that it required an externally set test and a portfolio of activities that was assessed by the provider. The two sources of information then had to be moderated. Finally, neither employers nor further education providers placed much weight on the qualification, so there was little point in students doing it or schools offering it.

The lessons for Australia from the Key Skills Qualification is that the report of achievement must have credibility with employers (and others), it must provide information that is not already conveyed by other achievement results, it should provide information about complex skills and the assessment regime must be relatively simple. This suggests that moderation may add a level of complexity that may not be acceptable in Australian jurisdictions.

Summary of standardised testing

A common feature of the four examples of standardised assessment is their focus on cognitive skills. The assessment methods used of the workplace skills elements of the Singapore ESS, that is, other than the basic skills, require further investigation. In the GSA, the assessment of teamwork used interpersonal understandings, a knowledge component of teamwork, as a proxy for the complete construct. This suggests that it is more difficult to measure constructs that have a more affective than cognitive character.

The difficulties experienced in the United Kingdom over the Key Skills qualification suggest that the employability skills should not be restricted to a narrow focus on basic skills and that the skills that are more difficult to assess, such as teamwork, should be included in any assessment of employability skills. The complexity of the Key Skills assessment, attributable to the moderation of internal and external assessments, should be avoided. Both internal and external assessments may occur, but they should be recognised as serving different purposes and reported separately. Engendering acceptance for reports of employability skills achievement by employers and other education providers will be important in ensuring that

students, parents, teachers and schools are prepared to invest in developing, teaching, assessing and reporting employability skills.

Table 1: Summary of generic employability skills included in instrumental assessment examples

Employability skills	Graduate Skills Assessment	PISA	Employability Skills Profiler	Key Skills Certificate (UK)
Communication	Written communication	Reading literacy	[Specific elements of all employability skills are assessed]	Written communication
Teamwork	Interpersonal understanding			
Problem solving	Problem solving	Problem solving Mathematical literacy		Application of number
Initiative and enterprise				
Planning and organising				
Self management				
Learning				
Technology skills				Information technology

Standardised testing is being used to assess skills having strong cognitive components, but with the exception of the ESP, not those of a more affective character. It is possible to assess attitudes, dispositions, motivations and values, but it is more difficult.

The PISA and GSA assessments meet reliability criteria. If the PISA assessment of problem solving is to be used as a model for the assessment of this and other employability skills, a pilot study to determine the precision of estimates of student achievement for different numbers of test items will be required. The three performance levels (with a fourth implied) of the PISA problem solving assessment provide useful information to and about learners. Similarly informative descriptors were developed for the PISA reading literacy and mathematics literacy assessments, with five and six levels identified, respectively (Thomson, Cresswell, & de Bortoli, 2004: 94 & 43). The identification of knowledge components of some of the ‘hard-to-measure’ skills such as the interpersonal understandings component of teamwork may provide a partial solution to the testing of the less tractable employability skills.

A central problem highlighted in the Employability Skills Profiler is the specificity of the skills assessed. General education is designed to prepare students for personal development, community engagement and workforce participation. With the possible exception of vocational education in senior secondary schooling, the workforce participation objective is necessarily broad, as students may move into very varied work roles. The notion that only specific elements of employability skills and not the employability skills themselves can be assessed as coherent constructs suggests that the current operationalisation of the skills

through their facets may be flawed. The construction of these skills as fixed traits is antithetical to the developmental role of general education.

The UK Key Skills Certificate provides a warning. It is apparent that the qualification did not add to what was known about students who were performing well academically and the test ceased to be used for these young people. The criticism that it was too narrowly focused and that the wider key skills were of greater interest to industry suggests that efforts need to be directed at assessing those constructs that clearly are difficult to assess.

Common assessment tasks

Common assessment tasks, because they are tasks not tests, typically allow for the assessment of a wider range of skills and practices than standardised tests, in a range of formats and settings that more closely approximate how people function in the wider world outside the classroom.

Whereas standardised tests are closely controlled at every stage of development, implementation and marking, the ‘commonality’ of common assessment tasks consists in their initial formulation and in the standards by which student performances will be assessed. Approaches to implementing the task are typically to some extent able to be decided by the teacher and perhaps the students. Implementation must, however, be within the common parameters and give students the opportunity to achieve highly against the standards.

Whereas the marking of standardised tests is often automated, the assessment of common assessment tasks requires expert judgment, typically by the teachers involved in the implementation of the task. The credibility of the grades awarded to students across (and even within) sites depends upon teacher–assessors having a shared understanding of the standards. A process of social moderation may help develop this shared understanding as well as provide a public basis for confidence in the grades awarded.

Whereas standardised tests have certain inbuilt limitations on the conditions of their implementation (in relation, for example, to time, location and the involvement of outside people), common assessment tasks have the potential to allow (or even require) long-term timeframes, cooperative participation of other students and community members, and ‘real-world’ approaches to getting things done.

Common assessment tasks example 1: Rich Tasks (Queensland New Basics)

The Rich Tasks, a component of the New Basics Framework, are an instance of such large-scale common assessment tasks. A Rich Task is the culmination of about three years’ work, in which the student masters the ‘targeted repertoires of practice’ for that task (the ‘cognitive and cultural, linguistic and social skills that need to be acquired developmentally in order to complete the Rich Task’ (Education Queensland, 2000: 39)). The Rich Tasks were designed to require students to engage in authentically real-world practices that require deep

understanding of disciplines of learning within a transdisciplinary approach. They involve the solving of substantive, real problems, and the undertaking of pragmatic social action.

Rich Tasks may be implemented in Years 1–9/10. Assessment and grading are conducted with reference to task-specific standards ('Desirable features'), and are subject to a statewide moderation process. A grade in a Rich Task constitutes a quality-assured indication of performance not in an individual discipline of learning but in a complex of knowledges, skills and attitudes required in a particular life role.

Common assessment tasks example 2: Queensland Comparable Assessment Tasks

A later development in Queensland than the Rich Tasks, the Queensland Comparable Assessment Tasks (QCATs)—a component of the Queensland Curriculum, Assessment and Reporting (QCAR) Framework—illustrate how common assessment tasks may also be applied to a discipline-based approach to curriculum, and implemented under more specified conditions. A QCAT assesses achievement in a designated subset of the 'Essential Learnings' (another component of QCAR) for a particular Key Learning Area. While the Essential Learnings are intended to encompass 'knowledge, skills and attributes that are required for complex, real-life challenges' (Queensland DET, 2005: 5), the QCATs themselves are more limited than Rich Tasks in the extent to which they are intended to embody large-scale life roles.

The conditions under which QCATs are implemented are more restricted than those for Rich Tasks and more liberal than those for standardised tests: the task itself is to be completed in under a few hours within a designated span of weeks, in Years 4, 6 and 9.

Each QCAT has task-specific standards, which relate to the standards associated with the Essential Learnings for a particular Key Learning Area. Possible models of moderation of QCATs are being investigated.

Teacher-generated tasks

'Performance-based assessment is a type of testing that calls for demonstration of understanding and skill in applied, procedural, or open-ended settings.' (Baker, O'Neil, & Linn, 1993: 1210). Performance assessment has been used in judging activities that do not normally leave an artefact that can be evaluated at a later time. Examples include gymnastics and dance, but the act of making an object is also a performance that can be evaluated independently of the object that is produced. Performance assessment has been extended to a wider range of activities, including science laboratory classes, medical diagnosis and building a brick wall. In each case, there is a product that can be evaluated, but employability skills are more likely to be observed in their execution during construction of the artefact than in an artefact that is produced. It is the act of producing rather than the product that is being judged.

Performance assessment differs from holistic teacher judgment. Performance assessment co-locates the act of judgment of employability skills with a specific performance by a student. It differs from the Common Assessment Tasks model in that the choice of tasks is made by the teacher, or possibly the learner, rather than using prescribed tasks.

Performance assessment is argued to be more authentic than pencil and paper alternatives, although the claim for authenticity depends upon the setting in which the performance is observed.

The validity of the assessment depends upon the context of the performance. In workplace assessment, the criterion practice is what is being performed and assessed, so the validity is high, provided that the criteria used to judge the performance do indeed reflect the desired characteristics of the intended practice. In school-based performance assessments, the validity might be judged by comparing the tasks and contexts with a criterion domain of practice.

A potential disadvantage of performance assessment is its reliability. If the performance is judged by a lone rater, no information is available about the reliability of the score. If multiple raters are used, their ratings can be compared, and if there is close agreement the score can be accepted as reliable. For high-stakes assessments, such as Olympic contests, multiple trained judges, strict criteria and scoring protocols are used. The complexity and cost of ensuring this level of reliability across schools and jurisdictions would be prohibitive and is not feasible. The reliability of performance judgments can be improved by the development of descriptive standards and consensus judgment by small panels within schools.

The assessment of achievement on a particular task will reflect the ability of the student, but it may be compromised by the severity of the person making the judgment and by the affordance of the task (Shavelson, Gao, & Baxter, 1993). The use of rubrics or other assessment tools may help to minimise the rater severity variability. The variability associated with the characteristics of the task may be minimised by requiring judgment across a range of tasks or by the use of prescribed tasks. (See the section on common assessment tasks.)

Alverno College

Alverno College is a recognised leader in the implementation of generic skills (abilities) in post-school education. Their eight abilities are embedded in the teaching and learning of discipline-based courses and feature in the assessment of those courses. Six performance levels for each of the abilities are recognised and students are required to achieve Level 4 on abilities assessments in all studies and Level 6 in their major studies. It is worth noting that the eight abilities are the focus of instruction and assessment within discipline-based courses

and that the abilities are assessed through the assignments that students undertake as part of their courses.

In addition to the assessment within courses, students undertake a number of interdisciplinary ‘integrative assignments’. These are assessed by panels of College faculty and community members who have undertaken training in the Alverno assessment framework. Alverno appears to enjoy considerable community support for such activities. The integrative assignments provide an opportunity to focus specifically upon the abilities. Alverno emphasises assessment *as learning* and students are encouraged to build ‘diagnostic digital portfolios’ of their assessments and to include reflections on their work and on the feedback they have received on their assignments. The portfolios are built from the outcomes of the performance assessments and are not themselves objects of assessment, although they could be assessed.

Coalition of Essential Schools

The Coalition of Essential Schools (CES) is a network of schools who share a commitment to educational reform built around ten common principles. These principles include a commitment to an assessment approach involving ‘multiple assessments based on performance of authentic tasks’—that is, performance assessment of ‘real world’ tasks and projects. A high school diploma is awarded following the demonstration of ‘mastery’ through an ‘exhibition’ of performances.

The [high school] diploma should be awarded upon a successful final demonstration of mastery for graduation—an ‘Exhibition.’ As the diploma is awarded when earned, the school's program proceeds with no strict age grading and with no system of credits earned by ‘time spent’ in class. The emphasis is on the students' demonstration that they can do important things.

(http://www.essentialschools.org/pub/ces_docs/about/about.html)

Members of the CES believe that individual teachers are more responsive to ‘economic signals and shifts in cultural values’ than central bureaucracies or testing organisations and are better able to interact with students in setting and interpreting standards for assessment (Rogers, McDonald, & Sizer, 1993).

The teacher-generated assessment model is used within the CES for routine assessment of courses. It could be used for employability skills if teachers judged them to be worthy assessment constructs.

Teacher-group judgment

Two variants of holistic teacher judgment are included in the assessment model framework (see Chapter 4), namely individual and collective teacher judgment. What is common to the two models is that teachers are asked to make ‘on-balance’ judgments of students’

demonstration of employability skills based on teachers' observations of students in both classroom and co-curricular activities.

Collective teacher judgment has been shown to work well in the school sector where teachers know students' attributes well through frequent and close observation (McCurry & Bryce, 1997, 2000). McCurry and Bryce established small panels of teachers and provided them with sufficient training in the key competencies to enable them to make consistent judgments of students' attainment of key competencies. This training and the observation of students, both in classroom-based and co-curricular activities, enabled teachers to make sufficiently consistent judgments to discriminate eight performance levels.

Holistic judgments by individual teachers are based on a summary of that teacher's observations of students in the classroom and co-curricular activities that the teacher has supervised. Clearly, the judgments made by individual teachers are limited compared with those made by groups of teachers. Individual teacher judgments are likely to reflect a smaller range of observed activities that are encompassed by a group of teachers and the group judgment is likely to even out less favourable and more favourable judgments of individual teachers.

Consistency of judgments within panels of teachers has been demonstrated. What has not been shown, at least in the context of employability skills, is that this consistency extends across school boundaries. All assessments, even multiple-choice tests, are informed by normative views of student performance. The range of difficulty of items or the standards that are prescribed in scoring rubrics are based, initially at least, on normative expectations of students. It is very likely that the norms that frame teachers' expectations within a school are informed by teachers' experiences in that school. The experiences of teachers in other schools will likely lead to different performance expectations. This means that students from different schools may be judged against different standards, and so individual achievement, assessed using this method, does not provide a basis for broad comparison.

Embedded development and assessment

One of the enduring problems, especially in the VET sector, has been misunderstanding of the relationship between alternative approaches to the development and assessment of employability skills. The view has been abroad that employability skills are either embedded (and therefore not overtly assessed), or delivered separately (and therefore explicitly assessed). In practice, there are other possibilities, and in particular, it is possible to have embedded delivery **and** explicit assessment. An attempt to clarify this dilemma is presented in Table 2.

Table 2: Possible arrangements for the delivery and assessment of employability skills

Assessment	Delivery	
	Embedded	Separate
Inferred	ESs are taught within existing subjects/courses. Performance is inferred from the achievement of unit/course objectives.	ESs are taught in separate subjects/modules. Since there is no other content, achievement is assumed. This is a non-assessment model.
Explicit	ESs are taught within existing subjects/courses. Performance is assessed explicitly using criteria or standards derived from descriptions of ES constructs.	ESs are taught in separate subjects/modules. Assessment of ES is undertaken within these modules.

It is possible to separate delivery and assessment strategies. In a study undertaken by the Queensland Department of Employment and Training, both integrated and separate delivery and assessment approaches were trialled (Queensland DET , 2004). Although the authors indicated a preference for separate delivery, they did report that a high proportion of the students who were enrolled in the discrete units ceased to attend classes. It seems that students did not value this approach. Other research, in the higher education sector, found that if students perceive the assessment to be authentic (and embedded in existing and valued learning) they are more inclined to use deep learning strategies (Gulikers, Bastiaens, Kirschner, & Kester, 2006). These findings suggest that the embedded or integrated delivery of employability skills is likely to enhance the learning of these skills.

If the teaching and learning of employability skills are embedded within existing subjects, it is possible to have either inferred assessment of them, or explicit assessment, and this can take several forms, including standardised testing and performance assessment. If the employability skills are delivered in separate subjects, there may be no assessment of them, that is, their delivery could be assumed to have achieved the purpose of informing learners about them, or they may be assessed explicitly in some of the ways outlined in this report.

A mapping of elements of the employability skills against the curriculum of existing subjects is undertaken to show that the employability skills are contained within the curriculum. The assumption is made that achieving those curriculum outcomes, demonstrated through existing assessment, can only occur if the mapped employability skill has been exercised. If this assumption is valid, it should be the case that students who master vocationally relevant content would demonstrate the mapped employability skills in practice. Surveys of employers, however, have not supported this hypothesised observation (AC Nielsen Research Services, 2000). It seems that explicit attention to employability skills in both teaching and assessment is required in order to achieve the outcomes desired for employability skills.

Authentic performance-based assessment

A study of key competencies assessment and reporting was undertaken in the VET sector by Curtis and Denton (2003). The research was undertaken in an electronics program at the Torrens Valley Institute of TAFE and focused on problem-solving skills.

An assessment tool was developed to scaffold instructors' judgments of student performance. The tool was based on a cognitive theory of problem solving (J. Bransford, Sherwood, Vye, & Rieser, 1986; J. D. Bransford & Stein, 1984). That model of problem solving posited five main processes, namely identification of the problem, definition of solution requirements, enacting a solution method, analysing the outcome and reflecting on the approach. Performance levels on indicators of each of the main processes were based on levels of the SOLO taxonomy (Biggs & Collis, 1982).

The assessment approach involved several steps.

- Existing assessment tasks in course modules were evaluated by staff members to identify opportunities for students to find and solve problems. Tasks were preferred if they provided scope for students to exercise all the problem-solving processes; that is, the tasks had to be challenging. The use of extant tasks is the basis of the claim for authenticity, as assessment tasks in VET programs are designed to be industry-relevant.
- A range of tasks were selected, some being relatively simple ones in which problems were provided through to tasks that were thought to provide a high level of challenge.
- Students were given a copy of the assessment tool so that they knew what the target processes were and what performance levels were being sought.
- Students undertook the assessment tasks and they were assessed for their substantive electronics content and the results were recorded as they normally were. However, students could choose to have the activity assessed for problem solving. This is an instance of assessment being embedded in routine teaching and learning, but also being explicitly of problem solving. When students submitted their work for problem solving assessment, they indicated on the assessment tool, evidence of their application of each of the component processes and of the level they believed they had achieved. The assessment form was submitted to their instructor at the same time as their work was submitted for routine assessment.

Instructors provided feedback to students about their problem solving performance. Where the instructor believed that the level of performance on an indicator was different from the student's perception, the feedback was directed at improving the student's understanding of the processes and their performance-level criteria.

The result of the problem-solving assessment was recorded separately from the routine assessment result. When students had been assessed on a skill on two occasions in different contexts, their achievement of that key competency was recorded.

A similar approach is applied to all key competencies (now employability skills).

Feedback from students suggested that a major outcome has been an improvement in students' knowledge of employability skills and their ability to describe in detail their approaches to solving problems. Consultations with employers indicated that they valued the attention that was being paid to developing these skills in students.

The method may have application in the schools sector. Whether tasks and their assessment criteria are set centrally or by individual teachers, an assessment tool that scaffolds judgments by teachers and learners could lead to improved understanding of what each of the employability skills means in practice. Articulating the performance-level criteria should help to inform students about desired achievement levels. The assessment is authentic to the extent that the tasks reflect real-world applications of the employability skills.

Portfolio construction

A portfolio is an assemblage of evidence of activities through which certain abilities are attested. The content of a portfolio can be quite varied. It could include: objective documentary evidence of achievement such as a certificate awarded following successful completion of a course; letters and testimonials of performance; photographs or videos of an individual undertaking a relevant activity; an individual journal of activities, perhaps with reflective comments.

The portfolio may be assembled and presented simply as a compendium of evidence or it could be subject to assessment as a body of evidence. In most cases reviewed, the emphasis was on the construction of the portfolio and it was not assessed. In such cases, if it is regarded as an act of assessment, it must be judged as an instance of self-assessment.

Portfolio assessment was recommended as the preferred method of assessing and reporting achievement of the employability skills across the schools, VET and higher education sectors (Allen Consulting Group & National Centre for Vocational Education Research, 2004). Although many other assessment and reporting options had been raised during consultations, the final report only dealt with portfolios. The reasons for this are clear as the criteria proposed to evaluate options were: effective, efficient, accountable, and transparent (Allen Consulting Group, 2004). The efficiency criterion seemed to have weighed heavily in their deliberations and portfolios were recommended as having the minimum impact on the need for professional development of teaching staff and the minimum assessment load. Neither specific portfolio templates, nor any particular methods for assessing or verifying portfolio content were suggested. Indeed a degree of latitude was also suggested in interpreting the employability skills themselves.

The review failed to acknowledge the literature on portfolio assessment. First, the construction of a portfolio is not an act of assessment. It produces an artefact that is available for assessment and the act of creating it may be a learning experience for the individual. The assessment of portfolios is a time-consuming task and that assessment is beset by low validity and reliability (Troper & Smith, 1997). Portfolios have low validity because the quality of a portfolio is a result of more than the target construct (demonstrated performance on an employability skill). The form of the portfolio also influences raters' judgments.

Portfolios do serve a useful purpose. They are repositories for detailed evidence of experience and performance. Alverno College encourages students to develop a 'diagnostic digital portfolio' in which students record the results of and their own reflections on other assessments, including externally judged integrative assessments.

There are many examples of portfolio templates and tools (Allen Consulting Group & National Centre for Vocational Education Research, 2004: 22–35).

Self-assessment

No stand-alone examples of self-assessment have been located, although this method of assessment has been advocated because of its purported benefits for student learning (Boud, 1995, 2002; Wiggins, 1998). Three of the cases described above include elements of self-assessment.

The construction of a portfolio involves self-assessment in that students are required to identify experiences through which they have developed or demonstrated target skills. In the case of portfolios designed as repositories of evidence of employability skills, students would normally be expected to make a case, perhaps through reflective comments, that the evidence they present does indeed reveal their achievement of employability skills. These actions require students to know what the employability skills entail, recognise these elements in what they have done and reflect on what they have done to argue a case based on their evidence. If the portfolio is assessed by a teacher, the quality of the evidence and the case built on it can be evaluated. Feedback to the student should help them to understand the performance standards by which their evidence is judged and should improve their self-assessment capability. Without the assessment of a recognised judge, the self-assessment is likely to carry little weight, irrespective of the value to the individual. Not all examples of portfolios used to accumulate evidence involve assessment or feedback to learners. Those cases in which a summative grade is given to a portfolio, without the opportunity for learners to revise their evidence, do little to develop learners' capacities for self-assessment.

The type of portfolio desired to support students' claims for their employability skills achievement is what Forster and Masters (1996, pp. 23-36) call a documentary portfolio. This type of portfolio is designed for assessment. Formal assessment of the portfolio may not be intended, but if the portfolio is to be presented as evidence in support of a claim for

employability skills achievement at, for example, an interview, a judgment will be made about the value of the evidence.

Self-assessment is a desired goal of the Alverno abilities-based curriculum (Loacker, 2000). The argument for self-assessment at Alverno is the same as that advanced elsewhere—that students' abilities to assess their own performance is an important facility for future learning. At Alverno, students assemble a portfolio of evidence to document what they have done in developing their 'abilities'.

In the embedded performance-based approach to generic skills assessment reported above (Curtis & Denton, 2003), self-assessment was a deliberate target for development. Students were given the assessment tool that their instructors would use in the final assessment and given similar information about how to use it. It was intended to guide students in their search for evidence of the generic skill processes and to help them assess the performance level they had demonstrated. The long-term objective of this approach was to develop the facility for self-assessment and, through this, for quality improvement. In Curtis and Denton (2003), the students' self-assessed grade was not recorded in the student records system.

Self-assessment is not being advocated as the basis for reporting student achievement of employability skills. However, if there is interest in preparing students for future learning and quality improvement, developing a capacity for self-assessment is desirable. Finding a way of having parallel assessment approaches, one involving assessment by a credible figure (e.g. a teacher) and the other involving self-assessment, with feedback between the systems, should provide credibility for reporting and build the capacity for learning.

References

- AC Nielsen Research Services. (2000). *Employer satisfaction with graduate skills. Research report*. Canberra: Department of Education, Training and Youth Affairs.
- Allen Consulting Group. (2004). *Development of a strategy to support the universal recognition and recording of employability skills: Issues paper*. Canberra: Allen Consulting Group.
- Allen Consulting Group, & National Centre for Vocational Education Research. (2004). *Development of a strategy to support the universal recognition and recording of employability skills. Directions Paper*. Melbourne: Allen Consulting Group.
- Australian Council for Educational Research. (2001). *Graduate Skills Assessment. Summary Report*. Canberra: DETYA, Higher Education Division.
- Australian Council for Educational Research. (2003). *Graduate Skills Assessment. Sample questions*. Melbourne: Australian Council for Educational Research.
- Baker, E. L., O'Neil, H. F., & Linn, R. L. (1993). Policy validity prospects for performance-based assessments. *American Psychologist*, 48(1210–1218).

- Biggs, J. B., & Collis, K. F. (1982). *Evaluating the quality of learning: the SOLO taxonomy (Structure of the Observed Learning Outcome)*. New York: Academic Press.
- Boud, D. (1995). *Enhancing learning through self assessment*. London: Kogan Page.
- Boud, D. (2002). Assessment: have we forgotten something? In Australian National Training Authority (Ed.), *Fresh thinking about learning and learners: a blue sky project* (pp. 42–45). Melbourne: Australian National Training Authority.
- Bransford, J., Sherwood, R., Vye, N., & Rieser, J. (1986). Teaching thinking and problem solving. *American Psychologist*, 41(10), 1078-1089.
- Bransford, J. D., & Stein, B. S. (1984). *The ideal problem solver: a guide for improving thinking, learning, and creativity*. New York: W H Freeman.
- Curtis, D. D., & Denton, R. (2003). *The authentic performance-based assessment of problem-solving*. Adelaide: NCVER.
- Forster, M., & Masters, G. N. (1996). *Portfolios*. Melbourne: ACER.
- Gable, R. K., & Wolf, M. B. (1993). *Instrument development in the affective domain. Measuring attitudes and values in corporate and school settings* (2nd ed.). Boston: Kluwer Academic Publishers.
- Gulikers, J. T. M., Bastiaens, T. J., Kirschner, P. A., & Kester, L. (2006). Relations between student perceptions of assessment authenticity, study approaches and learning outcome. *Studies in Educational Evaluation*, 32, 381–400.
- Hodgson, A., & Spours, K. (2000). *Qualifying for Success: Towards a framework of understanding*. London: Institute of Education.
- Loacker, G. (Ed.). (2000). *Self assessment at Alverno College*. Milwaukee, Wisconsin: Alverno College Institute.
- McCurry, D., & Bryce, J. (1997). *The school-based key competencies levels assessment project. Final report*. Canberra: DEETYA.
- McCurry, D., & Bryce, J. (2000). *Victorian Board of Studies. Key Competencies levels assessment trial. Working paper 2*. Melbourne: Victorian Curriculum and Assessment Authority.
- OECD. (2004). *Problem solving for tomorrow's world. First measures of cross-curricular competencies from PISA 2003*. Paris: OECD.
- Powell, R., Smith, R., & Reakes, A. (2003). *Basic skills and key skills: A review of international literature. Final report*. London: NFER.
- Queensland Department of Education and the Arts. (2004). The New Basics project. Retrieved 7 August, 2006, from <http://education.gov.qld.au/corporate/newbasics/index.html>
- Queensland Department of Education and the Arts. (2005). *Queensland curriculum assessment and reporting framework* (PDF document). Brisbane: Queensland Department of Education and the Arts.

- Queensland Department of Employment and Training. (2004). *Work futures: Employability skills and attributes*. Brisbane: Queensland Department of Employment and Training.
- Rogers, B., McDonald, J., & Sizer, T. (1993). Standards and school reform: Asking the essential questions. Retrieved 5 July, 2007, from http://www.essentialschools.org/cs/resources/view/ces_res/106
- Rust, J., & Golombok, S. (1999). *Modern psychometrics. The science of psychological assessment* (2nd ed.). London: Routledge.
- Shavelson, R. J., Gao, X., & Baxter, G. P. (1993). *Sampling variability of performance assessments* (CSE Technical Report No. 361). Los Angeles: University of California, CRESST.
- Thomson, S., Cresswell, J., & de Bortoli, L. (2004). *Facing the future: A focus on mathematical literacy among Australian 15-year-old students in PISA 2003*. Melbourne: ACER and OECD.
- Troper, J., & Smith, C. (1997). Workplace readiness portfolios. In H. F. O'Neil (Ed.), *Workforce readiness. Competencies and assessment* (pp. 357-382). Mahwah, NJ: Lawrence Erlbaum.
- Turner, D. (2002). *Employability skills development in the United Kingdom*. Adelaide: National Centre for Vocational Education Research.
- Wiggins, G. P. (1998). *Educative assessment. Designing assessment to inform and improve student performance*. San Francisco: Jossey-Bass.