

Teachers Making Contextualised, Group Judgements of Generic Skills and Dispositions

A Paper for the 32nd IAEA CONFERENCE

This paper outlines a project of the Victorian Curriculum and Assessment Authority, in collaboration with the Australian Council for Education Research that has been running since 2001. Different teachers of a student make independent assessments of the generic skills and dispositions of that student and these separate judgements are subsequently brought together into a single overall or whole-school report. The process is supported by specially developed computer software to synthesise the judgements of different teachers. The judgements made are generic and cross-curricular rather than subject-based.

More than 10 000 students in more than 100 schools have been assessed with the procedures developed in the project since 2001. The trials have shown that participating teachers are able to make global, impression judgements of generic skills and dispositions, and that the assessment process developed by the project is cost effective. The trials also showed a significant degree of agreement between different teachers about the individual student. The degree of agreement between teachers from across the curriculum made the production of an overall, collective or whole-school report both meaningful and feasible.

Participating schools have used this generic skills assessment and reporting model for a variety of purposes.

Doug McCurry
Senior Research Fellow
Australian Council for Education Research

Margaret Mackenzie
Manager, Vocational Education
Victorian Curriculum and Assessment
Authority

Introduction

This paper reports on the work of the Victorian Curriculum and Assessment Authority (VCAA) to implement an important national policy initiative in Australia, the Employability Skills Framework. The VCAA provides curriculum and assessment for the preparatory years of schooling to Year 12. It is a statutory body directly responsible to the Minister for Education and Training in Victoria and serves government and non-government schools. Victoria is one of eight states and territories in Australia. Each state and territory has its own state government and state education and training authorities. At the national level, Australia has a federal government with a federal Minister for Education and Training.

This work has evolved out of research undertaken by the Australian Council for Educational Research (ACER) into a model for assessing and reporting on student achievement of generic, or cross-curricular, skills. The development of the model outlined in this paper has proceeded in two parts: the testing of the feasibility of a skills schema targeting senior secondary students; and the testing of a methodology and accompanying software for assessment and reporting on this skills schema.

General education versus vocational education

The 1980s saw recession and substantial levels of unemployment in developed countries, followed by increasing globalisation of trade and an intensification of international economic competition. As a result of these pressures, a 'restructuring' of industry seemed an imperative in a number of developed countries in the late 1980s. As this economic reorganisation took place, an increasing desire for change in the vocational and general education systems became evident. The economic restructuring of the 1980s required changes in work and work practices, and the need for such changes became the impetus for an examination of the relationship between education and work in many countries.

In Australia, as in many other countries, there has been long and wide debate about the best ways to bridge the divide between general education and vocational education and training. For many years the schooling sector was seen to be responsible for the development of academic and cross-curricular skills of students in general education. The vocational education and training sector (VET) was seen to be responsible for preparing young people for the world of work. This preparation for the world of work was primarily achieved through training in specified skills relating to a specific industry.

VET in Schools

A merger between general and vocational education and training has occurred through a well established national system called VET in Schools. Approximately one third of senior secondary students (between the ages of 15 and 19 years) in Victoria choose to undertake a vocational training program as part of their senior secondary certificate. It is expected that they will develop skills to make them employable in specific industry areas but that they also will develop more general employability skills through their training program. VET in Schools has been extremely successful in delivering outcomes for students.

But both federal and state governments want all students in their senior years of schooling to develop employability skills, so the challenge for the schooling system is to find ways in which to comprehensively bridge the divide between general education and vocational education and training.

Key Competencies and Employability Skills

A number of reports since the early 1990's, commissioned by federal and state governments, have delivered important messages for policy development in the education and VET sectors. These reports have all emphasized the importance of the Key Competencies (as they are called in Australia) as identifying the underpinning skills necessary for a productive employee and therefore a productive workplace.

A distinctive feature of these Australian Key Competencies in comparison with similar work in other countries was the intention that these competencies be assessed and certified. While nothing has come of the original proposals for the assessment of Key Competencies in

general education other than to integrate them into curriculum, the VET sector has not only made reference to the Key Competencies in the nationally endorsed VET training specification documents, called ‘training packages’, it also claimed they were assessed as part of the assessment of the training package qualifications.

In 2001, the Commonwealth Department of Education Science and Training and Australian National Training Authority funded a project undertaken by the Australian Chamber of Commerce and Industry and the Business Council of Australia to analyse and report on current business requirements for ‘employability skills’ (ACCI and BCA 2002).

The Employability Skills project involved consultation with business and industry groups. The report of this consultation, *Employability Skills for the Future*, said that business and industry required a broader range of skills than the Key Competencies. The research found that business placed particular value on certain personal attributes that were not part of the original Key Competencies. Business indicated that personal attributes were as important as other employability skills, and that learning skills and self-management skills should also be included in the new framework.

Education authorities have been reluctant to endorse the set of personal attributes spelled out in the Employability Skills Framework because of the difficulty of assessing attributes. But there is general agreement that the skills sets articulated in the Employability Skills Framework are an advance on the original Key Competencies of the 1990’s. With the advent of the Employability Skills Framework, general education and vocational education and training have moved a step closer to a partnership, with a renewed emphasis on non-academic and cross-curricular skills.

Towards an integrated model

In Victoria, we are confident that we have a model for assessing and reporting student achievement of these non-academic skills that can be applied across the system. A collaboration between the VCAA and McCurry and Bryce of the Australian Council for Educational Research began in 2001 with the limited testing in schools of a methodology and software tool that assessed and reported on student achievement of the Key Competencies. Work continued on refining the performance level descriptors and the software over the next three years. With the release of the *Employability Skills for the Future* report, the VCAA asked the ACER consultants to investigate the feasibility of integrating the Key Competencies with the skill sets of the Employability Skills Framework.

The VCAA was mindful that this integrated skills construct had to be easily understood and acceptable to the schooling sector and therefore had to be described in terms that were not employment – context dependent. It was argued that with some modification, the Employability Skills construct can be understood in both a general education setting and a vocational education and training setting. The ACER consultants produced an integrated skills construct that is called Generic Employability Skills.

Table 1 The Employability Skills Framework – a summary

Personal attributes	
<ul style="list-style-type: none"> • loyalty • commitment • honesty and integrity • enthusiasm • reliability • personal presentation • commonsense • positive self-esteem • sense of humour • balanced attitude to work and home life • ability to deal with pressure • motivation • adaptability. 	
Key Skills	
Communication	skills that contribute to productive and harmonious relations between employees and customers
Team work	skills that contribute to productive working relationships and outcomes
Problem-solving	skills that contribute to productive outcomes
Initiative and enterprise	skills that contribute to innovative outcomes
Planning and organising	skills that contribute to long-term and short-term strategic planning
Self-management	skills that contribute to employee satisfaction and growth
Learning	skills that contribute to ongoing improvement and expansion in employee and company operations and outcomes
Technology	skills that contribute to effective execution of tasks.

In 2004 and early 2005, in consultation with schools, McCurry and Bryce developed for the VCAA a schema based on the above framework.

Table 2 Proposed VCAA Generic Employability Skills

Interdisciplinary skills	Personal skills	Dispositions
<p>Written Communication Writes accurately and conventionally writes clearly and coherently Uses formal and informal styles appropriately</p> <p>Oral Communication Speaks clearly and precisely Uses formal and informal speaking styles appropriately Responds in oral interchanges</p> <p>Logical reasoning Reasons systematically Employs logic Reasons quantitatively</p> <p>Interpretive reasoning Reasons critically Understands meanings Is sensitive to suggestions</p>	<p>Planning and organising Thinks ahead and anticipates possible problems Is systematic and practical Seeks organisational challenges and experiences</p> <p>Understanding and working with others Can read and understand others Adapts to and contributes to group processes Is sensitive to and supportive of others in a group</p>	<p>Initiative and enterprise Responds positively to challenges Sets and pursues own goals Is confident and adaptable</p> <p>Approach to Learning Has a positive attitude to learning Seeks opportunities to learn Reflects on own learning</p> <p>Approach to Technology Has a positive attitude to technology Can understand technology systems Seeks technology experiences and challenges</p>

McCurry and Bryce recommended that the assessment of personal attributes (proposed by the ACCI and BCA) is best approached by viewing such conceptions as more general than skills. While an attribute can involve skill, it can also involve personal characteristics. McCurry and Bryce argued that the term 'disposition' is more precise than 'attribute'. Ron Ritchhart in his book *Intellectual Character*, describes dispositions as 'acquired patterns of behavior that are under one's control and will' they are thus not 'automatically activated' (Ritchhart 2004). According to Ritchhart's conceptualisation, dispositions are 'coupled' with abilities. In the light of these distinctions it was proposed that the GES be understood as having a sub-set of constructs that are viewed as 'Dispositions'. These dispositions would be understood as a mixture of attitudes and aptitudes. On this basis it was proposed that the constructs of initiative and enterprise, learning skills ('Approach to learning') and technological skills ('Approach to technology') be viewed as dispositions.

The VCAA Whole-school Assessment trials for assessing and reporting student achievement of non-academic skills

Between 2002 and 2004 a total of 94 schools participated in a trial of software and teacher support materials for assessing and reporting student achievement of generic skills. These trials tested the proposition that reliable, global, whole-school judgments can be made about an individual student's achievement of the generic skills.

The model developed by the ACER consultants uses a custom-built software tool and process that generates a global judgement of student achievement derived from assessments provided by all the student's teachers. Students were assessed by as many of their teachers as possible. No special activities were required to be undertaken by the students to be assessed. Part of the trial was to test the extent to which generic skills are already embedded in existing curriculum and pedagogy. Towards the end of the semester participating teachers made a global judgement of the generic skills performance of the target students. This assessment took a teacher no more than 5 minutes per student. The assessment of teachers was entered into a piece of software for integrating different judgements and producing a recommended result overall.

In 2005, the VCAA and ACER conducted a trial for *Assessment and Reporting of Generic Employability Skills* which used the revised, integrated skills construct and assessment schema to coincide with the semester 2 reporting period for senior secondary students.

The trial participants comprised 28 secondary schools and 3 TAFE (Technical and Further Education) institutes. The assessment trials undertaken by the VCAA were based on the model of Whole-school Assessment developed by McCurry and Bryce of the Australian Council for Educational Research. (See Appendix for the results of this work.)

What is Whole-school Assessment of Generic Skills

Whole-school assessment is the assessment of generic skills by groups of teachers contributing to an overall or cross-curricular report on a student. Apart from tertiary entrance ranks or grade-point averages, almost all educational assessment is done by particular teachers in particular subject areas. Subject-based Individual Teacher assessment (SIT) is so predominant that it seems difficult to conceive of any other kind of teacher assessment. The alternative to Subject-based Individual Teacher assessment is Group Generic assessment (GG). Group Generic assessment is a single overall assessment of generic or cross-curricular skills made by groups of teachers.

GG assessment could be a matter of groups of teachers making judgements about the generic skills demonstrated in more or less specific tasks. Such an approach to GG assessment would involve new and different assessment tasks being added to educational programs, and teachers would have to add another kind of assessment to their current workload. This form of GG assessment is not feasible or cost-effective. The most realistic and practical form of GG assessment is made in normal programs, and it involves global impression judgements rather than judgements about more or less specific tasks. McCurry and Bryce have called this kind of assessment Whole-school Assessment (WSA).

WSA Assessment is:

- generic and cross-curricular;
- made by groups of teachers rather than individual teachers;
- based on global impression judgements; and
- made in normal programs rather than in special generic assessment tasks.

Global Impression Judgements of Generic Skills

The judgements made in the WSA process are concerned with that which is general to education (and might be properly called cross-curricular skills) rather than the knowledge and skills that are more or less specific to individual subjects and subject-based assessments.

It is crucial that the judgements made by individual subject teachers in the WSA process are general rather than subject specific. It is understood that teachers will gain most (but not all) of their knowledge of students from subject classes, but teachers in WSA assessment are not being asked to make judgements about that which is specific to their subject area. The WSA assessments are **not** about students' performances in particular subject areas. They are judgements about what the particular teacher takes to be general about the abilities or the performances of a student. These judgements are broad or global inferences about what is taken to be typical of the students in most subject areas, and other activities.

If the generic skills assessed in the WSA process are to be general, they must take into account as much of the school activity of a student as possible. In practical terms, an overall judgement that arises from a range of teacher perspectives can be seen as escaping the limitations of any individual teacher's perspective, and, as a result, increasing the reliability of the assessment.

Teachers of different subject areas have different perspectives on students, and differences in their judgements of students' generic skills are to be expected. But such differences should not result from teachers thinking they are assessing different things or because they assume that what is true of their subject or their personal interaction with the student is typical or general. For instance, a student who has a gift for music may be a very positive member of a school orchestra, but this does not mean that the music teacher who organises the orchestra should automatically assume or claim that such activity amounts to a high level of **Understanding and working with others** in general. The performance of such a student is the basis for claiming a high level if the music teacher considers it a general characteristic of the student. It is possible that although the music teacher observes that student working well with others in the orchestra, s/he is doubtful about the student's ability to understand and work well with others in general, and so does not believe the student is typically at a high level in terms of understanding and working with others in general. Other examples of the difference between context or subject-specific performances that would be recognised in subject reports and the typical or general performances that are assessed in WSA levels might easily be elaborated. The crucial point for the WSA assessment is that these generic skills judgements are about what the teacher infers from what s/he observes about the general rather than the subject specific.

The Basis of the Global Impression Judgements in Whole-school Assessment

Towards the end of a semester teachers participating in a Whole-school Assessment are asked to reflect formally on and record their judgements on an 8 or 9 point scale (divided into 3 levels) in an Excel spreadsheet. It is presumed that contact in subject classes is the basis of this generic skills assessment, but other information gathered from co- and extra-curricular activities and work placements can be taken into account. The assumption behind the WSA process is that the judgements made in this assessment are stage-related, in that they are made explicitly about students in specified grade levels at a point in time. The judgements are made on the basis of participating teachers' knowledge and experience of students at this level and they are based on what teachers know and expect of students at this grade level.

The WSA assessments are based on the recognition that some students stand out, either because they exceed or do not meet usual and reasonable expectations. Students who do not meet the expected level of performance are judged to be below Level 2, and those who clearly

distinguish themselves as high level performers are at Level 3. These suggestions imply that most students will fall in Level 2 on a particular skill because they meet expectations for the level but are not clearly distinguished as high level performers. The distinction between Levels 2 and 3 is the difference between the expected achievement level for students and some clear signs of above expected achievement.

The Whole-school Assessment Process and Software

The whole-school Assessment Process and Software are built around the role of an Overall Assessor who uses the WSA software to review the global impression judgements made by different teachers. The software presents the assessments made by different teachers of a generic skill of a student on the following screen of the software.

Generic Skill Assessment

4.1 **Generic Skill Assessment**

GES1 GES2 GES3 GES4 GES5 GES6 GES7 GES8 GES9 GES10

Written communication

Student: Doug McCurry Result: 3 To 5
Assessor: Margaret Forster

Overall Range

Teacher	Teacher Mean	Grade
Geoff Masters	4.00	3
John Ainley	5.00	4
Margaret Forster	6.00	5

Enter your Assessment for this student at this time point

Lower: 3 Upper: 5

Accept Suggested Grades

This Skill and Student
 All Skills for this Student
 All Skills for all Students

Return

The software suggests a possible outcome for the student from the teacher input. The Overall Assessor reviews the input from the teachers and the suggestion of the software and decides to either confirm or change the proposed assessment. The Overall Assessor may occasionally need to consult with teachers over discrepant assessments

Making a Whole-school Assessment is dependent on having a relational database that can take input from different teachers and present it efficiently to an Overall Assessor. As well as being able to produce reports for individual students, the WSA software produces a wide range of statistical analysis of the skills of the students and the assessments of the teachers.

Data from the whole-school assessment can be used for:

- tracking students;
- monitoring over time;
- assessing different kinds of performance;
- focusing on issues (such as using technology or initiative);
- counseling students;
- reporting to students and parents; and

- reporting to others.

More than 10, 000 students in more than 100 schools have been assessed with the procedures developed in the whole-school assessment projects since 2001. The trials have shown that participating teachers are able to make global, impression judgements of generic skills and dispositions, and that the assessment process developed by the project is cost effective. The trials also showed a significant degree of agreement between different teachers about the individual student. The degree of agreement between teachers from across the curriculum made the production of an overall, collective or whole-school report both meaningful and feasible.

What have the VCAA trials of Whole-school Assessment shown us?

- The Generic Employability Skills construct and methodology have received broad endorsement from teachers.
- There is a significant degree of agreement of teachers in the assessments of students. The degree of agreement between teachers from across the curriculum made the production of an overall, collective or whole-school assessment both meaningful and feasible.
- There is a willingness of staff in school settings to embrace the model – both its principles and its practice.
- The benefits for students are well known and widely acknowledged in schools.
- The software as the primary tool of methodology is critical to the success of the assessment and reporting model.
- Individual schools/providers have used the model for achieving individual ends.
- The model has demonstrated the ability to be adapted to a variety of contexts and for a variety of purposes.
- The model has the capacity to stimulate pedagogical review and reinvigoration of school based curriculum development.

Conclusion

In a number of respects the results of the Whole-school Assessments trials are surprising and remarkable. The results show a significant degree of agreement between teachers of different parts of the curriculum in their assessments of the generic skills of the same student. The results are psychometrically sound, and they can carry useful information. The trials show that teachers can distinguish what they take to be general about a student from what they take to be subject-specific or particular to their relationship with the student. The trials also show that teachers see and know a good deal more about students than is included in subject assessments, and that this understanding can be drawn on in a cross-curricular assessment. The trial results suggest that the performance of students across different subject areas shows less variation than we might expect or assume.

The teachers within a school as a group seem to discriminate quite consistently among their students, and the trial results show that generic skills can be usefully assessed at provider level without a large investment of time or resources. The Whole-school Assessment procedures were judged by participants to be efficient and cost-effective. Teachers were able to make assessments of a student in two or three minutes using the method developed in the trial, and the procedures give the basis for producing integrated, Whole-schools Assessments efficiently.

One aspect of the Whole-school Assessment process that merits a good deal more research is the capacity of the model to generate curriculum debate in a school and to be a lever for change. Schools participating in the trials have consistently reported on the positive effects for curriculum development and review and review of reporting practices. The VCAA plans work of this nature in the near future.

Other areas of research and testing using this model will be undertaken with employer groups and TAFE providers of senior secondary programs. The VCAA is interested to evaluate the efficacy of the model in non-school settings.

Appendix

The Results of the First Whole-school Assessment Trial

Table I, II and III below present the results of the first whole-school assessment trial. Trial participants were interviewed after the assessment and they generally reported that they found the assessment framework was clear, they were able to use it with little difficulty, and they were able to make the global, impression judgements envisaged by the project. Table I shows that individual teachers spread the students over the 8 point score range and discriminated quite significantly amongst them.

Table I Frequencies of all assessment of all generic skills

Score	Percent	Cum. %	Percent	Cum. %	Percent	Cum. %
	GS1 Information		GS2a Oral		GS2b Writing	
1	3.5	3.5	1.4	1.5	3.5	3.6
2	3.0	6.6	1.4	2.9	2.3	5.9
3	18.9	25.7	15.8	19.0	19.3	25.6
4	12.8	38.7	12.8	32.0	12.7	38.6
5	31.1	70.0	32.7	65.2	30.0	69.3
6	11.5	81.6	14.0	79.4	12.7	82.2
7	16.2	97.9	18.6	98.3	15.6	98.1
8	2.1	100.0	1.7	100.0	1.8	100.0
Total	99.1		98.4		98.0	
	GS3 Mathematics		GS4 Cultural U		GS5 Solving Probs	
1	1.6	2.9	2.3	2.7	4.2	4.3
2	1.8	6.3	1.4	4.3	3.3	7.7
3	13.2	31.2	16.2	23.5	17.8	25.8
4	6.3	43.2	10.9	36.3	12.9	39.0
5	17.1	75.5	29.2	70.8	27.2	66.7
6	5.8	86.4	10.7	83.4	12.3	79.2
7	5.9	97.6	12.6	98.3	17.6	97.1
8	1.3	100.0	1.4	100.0	2.8	100.0
Total	52.8		84.7		98.2	
	GS6 Technology		GS7 Planning		GS8 Teamwork	
1	2.1	3.0	3.5	3.6	2.7	2.8
2	1.4	5.1	3.2	6.8	1.9	4.7
3	14.3	25.8	18.5	25.6	16.3	21.6
4	8.0	37.4	14.8	40.7	10.5	32.3
5	26.1	74.9	28.2	69.4	30.0	63.3
6	7.9	86.3	11.6	81.2	12.9	76.6
7	8.6	98.7	15.6	97.1	20.2	97.4
8	0.9	100.0	2.9	100.0	2.5	100.0
Total	69.4		98.1		97.1	

Table II shows the degree of agreement between different teachers assessing the same student. A difference of three score points (1 to 4, 5 to 8 etc.) on an 8 or 10 point scale is commonly judged to be unacceptably discrepant in a structured marking procedure dealing with a single piece of work. Marking procedures usually require some kind of resolution of this discrepancy. Such two marker discrepancies in controlled marking of a single piece of work should be under 10% and are good if they are under 5% on an 8 to 10 point scale.

Table II Percentage agreement in pairwise comparison of generic skills assessments

	GS1 Information			GS2a Oral			GS2b Writing		
	Freq	%	Cumul.	Freq	%	Cumul.	Freq	%	Cumul.
0	668	31.4	31.4	598	28.6	28.6	658	31.8	31.8
1	701	33.0	64.4	711	34.0	62.6	688	33.2	65.0
2	544	25.6	90.0	573	27.4	90.0	510	24.6	89.7
3	146	6.9	96.8	136	6.5	96.5	149	7.2	96.9
4	55	2.6	99.4	70	3.3	99.8	57	2.8	99.6
5	9	0.4	99.9	4	0.2	100.0	5	0.2	99.9
6	3	0.1	100.0				3	0.1	100
Total	2126			2092		100	2070		
Disc		10.0			10.0			10.3	
	GS3 Mathematics			GS4 Cultural U			GS5 Solving Problems		
	Freq	%	Cumul.	Freq	%	Cumul.	Freq	%	Cumul.
0	168	29.6	29.6	454	28.7	28.7	571	27.4	27.4
1	188	33.2	62.8	540	34.1	62.8	676	32.5	59.9
2	156	27.5	90.3	418	26.4	89.3	538	25.8	85.7
3	45	7.9	98.2	103	6.5	95.8	188	9.0	94.7
4	9	1.6	99.8	54	3.4	99.2	85	4.1	98.8
5	1	0.2	100.0	8	0.5	99.7	20	1.0	99.8
6				5	0.3	100	5	0.2	100
Total	567			1582	100		2083		
Disc		9.7			10.7			14.3	
	GS6 Technology U			GS7 Planning			GS8 Teamwork		
	Freq	%	Cumul.	Freq	%	Cumul.	Freq	%	Cumul.
0	298	29.9	29.9	585	28.2	28.2	546	27.2	27.2
1	300	30.1	60.0	712	34.3	62.5	676	33.6	60.8
2	298	29.9	90.0	542	26.1	88.6	555	27.6	88.4
3	68	6.8	96.8	162	7.8	96.4	141	7.0	95.4
4	26	2.6	99.4	66	3.2	99.6	74	3.7	99.1
5	2	0.2	99.6	8	0.4	100.0	14	0.7	99.8
6	4	0.4	100.0	1	0.0	100.0	5	0.2	100.0
Total	996			2076			2011		
Disc		10.0			11.4			11.6	
Disc average	10.9%								

Table II presents the agreement between all pairs of teachers assessing the same student using the Whole-school Assessment procedures. Rows 0 to 6 show the percentage of students results where there were no differences (row 0), 1 point difference (row 1), and so on. A difference of 1 is the difference between scores of 3 and 4 or 7 and 8. A difference of 2 is the difference between scores of 3 and 5 or 6 and 8. The last row for each generic skill in the table shows the discrepancies of more than 2 points difference. The average for adjacent grades was 55.8%, and the average for grades separated by a difference of 1 or 2 is 88.8% in these pairwise comparisons. These figures give an average pairwise discrepancy rate of 11.2% for these global impression judgements of generic skills. Table III is a summary of the 2 and 3 point differences from Table II.

Table III Percentage of students given a difference of 3 or less for pairs of teachers

	GS1	GS 2a	GS2b	GS3	GS4	GS5	GS6	GS7	GS8
% of no difference	31.4	28.6	31.8	29.6	28.7	27.4	29.9	28.2	27.2
% at or below a difference of 1	64.4	62.6	65.0	62.8	62.8	59.9	60.0	62.5	60.8
% at or below a difference of 2	90.0	90.0	89.7	90.3	89.3	85.7	90.0	88.6	88.4
Average difference of 1 or less			62.3%						
Average difference of 2 or less			89.1%						